

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 04:52:59 ; Search time 8109 Seconds
(without alignments)
11037.534 Million cell updates/sec

Title: US-10-015-387A-219

Perfect score: 2065

Sequence: 1 gtgaatgagggttgatg.....attaatatgttaaaaaa 2065

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.hcg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2065	100.0	2065	6	AX454630	Sequence
2	2065	100.0	2065	6	AX491108	Sequence
3	2065	100.0	2065	6	AX697151	Sequence
4	2065	100.0	2065	9	AX358527	Sequence
5	1318.4	63.8	117854	9	HSJ85A10	Human DNA
6	873	42.3	926	9	BC047076	Homo sapi
7	758.8	36.7	248469	2	AC015860	Homo sapi
8	635.6	29.8	183553	2	AC012335	Homo sapi
9	614	29.7	614	6	AX317844	Sequence
10	614	29.7	614	6	AX419522	Sequence
11	609.4	29.5	611	6	AX498264	Sequence
12	608	29.4	107500	10	AL929097	Mouse DNA
13	608	29.4	222104	10	AL929254	Mouse DNA
14	608	29.4	256869	2	AC117551	Mus muscu
15	594.6	28.8	110000	2	AC142108	Continuation (3 of
16	571.4	27.7	222922	2	AC106497	Rattus no
17	571.4	27.7	245485	2	AC135745	Rattus no
18	520.8	25.2	1286	6	AX402346	Sequence
19	514.2	24.9	808	10	AY134663	Mus muscu
20	492.2	23.8	183553	2	AC012335	Homo sapi
21	477.8	23.1	493	6	AX414692	Sequence
22	477.8	23.1	493	6	BD110245	EST and e
23	455.4	22.1	248469	2	AC015860	Homo sapi
24	375.2	18.2	1805	5	BC056851	Xenopus l
25	292.4	14.2	1810	10	BC055682	Mus muscu
26	281.2	14.1	2423	9	BC035789	Homo sapi
27	289.8	14.0	2484	6	AX018985	Sequence
28	289.8	14.0	2484	6	BD135080	Cerebelli
29	289.8	14.0	2750	6	AX463952	Sequence
30	289.8	14.0	2750	9	AV359074	Homo sapi
31	257.4	12.5	1524	9	HUMCERA	Human prece
32	254.8	12.3	2056	10	BC055730	Mus muscu
33	253.6	12.3	1266	5	BC054217	Xenopus l
34	239.2	11.6	2430	10	BC051956	Mus muscu
35	154.4	7.5	3336	9	AK125422	Homo sapi
36	152.6	7.4	285	6	AX913594	Sequence
37	152.6	7.4	285	6	BD049127	Sequence
38	151.6	7.3	265	10	S76975	Mus sp. cer
39	145.6	7.1	167232	5	AL929379	Zebrafish
40	145.2	7.0	181602	9	AC011208	Homo sapi
41	134.2	6.5	234942	10	AC115783	Mus muscu
42	132.8	6.4	165651	9	AC044798	Homo sapi
43	132.8	6.4	185418	2	AC012311	Homo sapi
44	132.8	6.4	192347	9	AC007614	Homo sapi
45	130	6.3	1384	6	BD193129	207 human

ALIGNMENTS

RESULT 1
AX454630
LOCUS
DEFINITION
Accession
Accession
Version
Keywords
Source
Organism
Reference
Authors

AX454630 2065 bp DNA linear PAT 06-JUL-2002
Sequence 215 from Patent WO0208284.
AX454630
AX454630.1 GI:21713939

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa;

Mammalia; Eutheria;

Primates; Catarrhini;

Hominoidea; Homo.

1

Baker, K.P., Ferrara, N.,

Gerber, H., Gerritsen, M.E.,

Goddard, A.,

Godowski, P.J., Gurney, A.L.,

Hillan, K.J., Marsters, S.A.,

Pan, J., Paoni, N.F., Stephan, J.P.,

Watanabe, C.K., Williams, P.M.,

Wood, W.I.

TITLE		and Ye, W.		Compositions and methods for the diagnosis and treatment of disorders involving angiogenesis	
JOURNAL		Patent: WO 0208284-A 215 31-JAN-2002;		Genentech, Inc. (US) ; Baker, Kevin P. (US) ; Ferrara, Napoleone (US) ; Gerber, Hanspeter (US) ; Grietsen, Mary E. (US) ; Goddard, Audrey (US) ; Godowski, Paul J. (US) ; Gurney, Austin L. (US) ; Hillan, Kenneth J. (US) ; Marsters, Scott A. (US) ; Pan, James (US) ; Praoni, Nicholas F. (US) ; Stephan, Jean-Philippe F. (US) ; Watanabe, Colin K. (US) ; Williams, P. Mickey (US) ; Wood, William I. (US)	
FEATURES		Location/Qualifiers			
source		1..2065			
		/organism="Homo sapiens"			
		/mol_type="unassigned DNA"			
		/db_xref="taxon:9606"			
ORIGIN		Query Match		Score 2065; DB 6; Length 2065;	
		Best Local Similarity 100.0%; Pred. No. 0;		Matches 2065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GTGAATGTGAGGGTTTGATGACTTTTCAGATGCTTAGGAACACAGAGTGGGTGCAGGGGCC	60		
DB	1	GTGAATGTGAGGGTTTGATGACTTTTCAGATGCTTAGGAACACAGAGTGGGTGCAGGGGCC	60		
QY	61	CAGGACGGCTGATTTCTTGGCGGAGGAGTAGGGTAAAGGTTCTGCATCAGCTCCTT	120		
DB	61	CAGGACGGCTGATTTCTTGGCGGAGGAGTAGGGTAAAGGTTCTGCATCAGCTCCTT	120		
QY	121	AAAGGACAAAGGTAAACAGACGACGAGAGCTCGAGGGGAGCTTTGACTTCAAGCCA	180		
DB	121	AAAGGACAAAGGTAAACAGACGACGAGAGCTCGAGGGGAGCTTTGACTTCAAGCCA	180		
QY	181	CAGAAATGGTGAAGTGTGCGCGCGCGCGCTCGCTCGAGCGCTCTGCACCTA	240		
DB	181	CAGAAATGGTGAAGTGTGCGCGCGCGCGCTCGCTCGAGCGCTCTGCACCTA	240		
QY	241	GCCGTAGCACTTCCCGAGACCGGGATCCCGGGTAGGAGCGACGCGGCGAGCAC	300		
DB	241	GCCGTAGCACTTCCCGAGACCGGGATCCCGGGTAGGAGCGACGCGGCGAGCAC	300		
QY	301	AGCGCAGCGCGCTCGCGTGCACACAGCGCTCACCANTGGCTCCGCGGCGCGGCGCTG	360		
DB	301	AGCGCAGCGCGCTCGCGTGCACACAGCGCTCACCANTGGCTCCGCGGCGCGGCGCTG	360		
QY	361	TCCGCGGTCCGCGCGCTGCTGCTCTCCTCAGCTGCGCGGCTGCGCGTCTGGGCAAG	420		
DB	361	TCCGCGGTCCGCGCGCTGCTGCTCTCCTCAGCTGCGCGGCTGCGCGTCTGGGCAAG	420		
QY	421	AACGACAGGAGCCCATCGTGTGAGGGCAAGTGTCTGGTGTGCGACTCGAACCCG	480		
DB	421	AACGACAGGAGCCCATCGTGTGAGGGCAAGTGTCTGGTGTGCGACTCGAACCCG	480		
QY	481	GCCAGGACTCCAAAGGGTCTCTTCTCTCCCGCTGGGGATATCGGTCCGCGGCGCCAA	540		
DB	481	GCCAGGACTCCAAAGGGTCTCTTCTCTCCCGCTGGGGATATCGGTCCGCGGCGCCAA	540		
QY	541	TCCAAAGTTCGCTTCTCGCGGTGCGGAGCAACACAGGCCATCCGAGATAGCAAC	600		
DB	541	TCCAAAGTTCGCTTCTCGCGGTGCGGAGCAACACAGGCCATCCGAGATAGCAAC	600		
QY	601	AAGACGCGCATCATTTACTTCGATCAGATCTCTGGTGAATGGGTAATTTTTCACATG	660		
DB	601	AAGACGCGCATCATTTACTTCGATCAGATCTCTGGTGAATGGGTAATTTTTCACATG	660		
QY	661	GAGTCTGCTTTGTAGCACCAGAAAGAAATTTACAGTTTCAGTTTTCAGTGAATAA	720		
DB	661	GAGTCTGCTTTGTAGCACCAGAAAGAAATTTACAGTTTTCAGTTTTCAGTGAATAA	720		
QY	721	GTCTACAGAGCCAACTATCCAGTTAACTTGATGTTAAATGGAAAAACAGTAATCT	780		
DB	721	GTCTACAGAGCCAACTATCCAGTTAACTTGATGTTAAATGGAAAAACAGTAATCT	780		

QY	781	GCCTTTGCGGGGACAAAGATGTTTACTCGTAGAGCTGCCACAAATGGTGTCTGCTCTAC	840
DB	781	GCCTTTGCGGGGACAAAGATGTTTACTCGTAGAGCTGCCACAAATGGTGTCTGCTCTAC	840
QY	841	CTAGATAAAGAGGATAAGGTTTACCTAAACCTGGAGAAAGTAATTTGTTGGAGGCTGG	900
DB	841	CTAGATAAAGAGGATAAGGTTTACCTAAACCTGGAGAAAGTAATTTGTTGGAGGCTGG	900
QY	901	CAGTATTCACGTTTCTTGGCTTTCTGGTGTTCCTCCCTATAGGATCAATTTCTCATGA	960
DB	901	CAGTATTCACGTTTCTTGGCTTTCTGGTGTTCCTCCCTATAGGATCAATTTCTCATGA	960
QY	961	TGTTTCATCCAGGTAGGATGACCCACCTCCCTGAGTTATTGGAAAGTCAATTTTTCATCAT	1020
DB	961	TGTTTCATCCAGGTAGGATGACCCACCTCCCTGAGTTATTGGAAAGTCAATTTTTCATCAT	1020
QY	1021	TGGATGTAGTCTTTTATTGGTTCCTCATGGGTGGATATGGAATCTTAAGGATTTCTAGCT	1080
DB	1021	TGGATGTAGTCTTTTATTGGTTCCTCATGGGTGGATATGGAATCTTAAGGATTTCTAGCT	1080
QY	1081	GTCTGAACCAATACAAATTTTTCACAGATTTTGTGTGTCTGTTCAGTATATTTTGA	1140
DB	1081	GTCTGAACCAATACAAATTTTTCACAGATTTTGTGTGTCTGTTCAGTATATTTTGA	1140
QY	1141	TTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTAACTCAAAAGCTGTCTGCAAG	1200
DB	1141	TTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTAACTCAAAAGCTGTCTGCAAG	1200
QY	1201	ACTTATTCGAAATTTCAATTTCTGGGATTTACTGAAATAGTTTACAGATGTTGAAATTTT	1260
DB	1201	ACTTATTCGAAATTTCAATTTCTGGGATTTACTGAAATAGTTTACAGATGTTGAAATTTT	1260
QY	1261	TGTTTAGTCTTAAAGACTGGCAACAGCTCTAAGGATTAGAAACTCTAAAGTCTGAC	1320
DB	1261	TGTTTAGTCTTAAAGACTGGCAACAGCTCTAAGGATTAGAAACTCTAAAGTCTGAC	1320
QY	1321	TTCAATCAACCGTTAGTGTGATATCTGCCAAAGAACTGTATACTGTGTTTAAATATTTG	1380
DB	1321	TTCAATCAACCGTTAGTGTGATATCTGCCAAAGAACTGTATACTGTGTTTAAATATTTG	1380
QY	1381	ATATTTGTTTATTTCTTTGGAATTTAGTTTGTGTTTCTGTGTTTAAATATTTG	1440
DB	1381	ATATTTGTTTATTTCTTTGGAATTTAGTTTGTGTTTCTGTGTTTAAATATTTG	1440
QY	1441	TTTTTTCAGTAACTGGTATTTATTTCTCTTAAATAAGTAAATGATGCTGTTGCCAC	1500
DB	1441	TTTTTTCAGTAACTGGTATTTATTTCTCTTAAATAAGTAAATGATGCTGTTGCCAC	1500
QY	1501	AAATTTTACCTTGACTAGATATCTGACATGACTTCTCTCAAAAAAAGAAATGCTTCA	1560
DB	1501	AAATTTTACCTTGACTAGATATCTGACATGACTTCTCTCAAAAAAAGAAATGCTTCA	1560
QY	1561	TAGTTGTATTTTAAATTTGTAATGTAAGAGTCAATTTTCCCAAGTATATTTTCTAAGA	1620
DB	1561	TAGTTGTATTTTAAATTTGTAATGTAAGAGTCAATTTTCCCAAGTATATTTTCTAAGA	1620
QY	1621	AGAAATAGATCATATAATCTGCAAGCAAAAAAGTCTCTTACCCAAATCTAAGTGTCTCA	1680
DB	1621	AGAAATAGATCATATAATCTGCAAGCAAAAAAGTCTCTTACCCAAATCTAAGTGTCTCA	1680
QY	1681	ATCCCTGAGGCTCAGCAAAAACAGCTCCCTCCGAGGGAATCTTATACTTTATTTGCTCAA	1740
DB	1681	ATCCCTGAGGCTCAGCAAAAACAGCTCCCTCCGAGGGAATCTTATACTTTATTTGCTCAA	1740
QY	1741	CTTTTAAATTAATGATTAATAACCACTTTTAAATAAACCTTAAGGTTTTTTTTTTTTC	1800
DB	1741	CTTTTAAATTAATGATTAATAACCACTTTTAAATAAACCTTAAGGTTTTTTTTTTTTC	1800
QY	1801	CGTAGACATGACCACTTTTAACTGGTGGGATGCTGTGTTTCTTAATATACCTAT	1860
DB	1801	CGTAGACATGACCACTTTTAACTGGTGGGATGCTGTGTTTCTTAATATACCTAT	1860


```
Db 1441 TTTTTCAGTAAGTGGTATATGTTTCTCTTTAAATAAGTAAGTGGCTTGCCAC 1500
QY 1501 AAATTACCTTGACATACGATATCATCGATGACATCTCTCAAAAAGAAAGTCTTCA 1560
Db 1501 AAATTACCTTGACATACGATATCATCGATGACATCTCTCAAAAAGAAAGTCTTCA 1560
QY 1561 TAGTGTATTTAAATGTAATGTGAAGAGTCATATTTTCCAAGTATATTTTCTAAGA 1620
Db 1561 TAGTGTATTTAAATGTAATGTGAAGAGTCATATTTTCCAAGTATATTTTCTAAGA 1620
QY 1621 AGAAGAATAGATCATATAATCTGACAGGAAAGTGTCTTACCCAAATCTTAAGTGCTCA 1680
Db 1621 AGAAGAATAGATCATATAATCTGACAGGAAAGTGTCTTACCCAAATCTTAAGTGCTCA 1680
QY 1681 ATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGGAAATCTTATATCTTTATGCTCAA 1740
Db 1681 ATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGGAAATCTTATATCTTTATGCTCAA 1740
QY 1741 CTTTAAATTAAGTATGATAATAACCACTTTATTAATAAACCTTAAGTGTCTTCTTTC 1800
Db 1741 CTTTAAATTAAGTATGATAATAACCACTTTATTAATAAACCTTAAGTGTCTTCTTTC 1800
QY 1801 CGTAGACATGACCATTTATTAATGATGCTGGTGGGATGCTGTGTTCTAAATTAACCTAT 1860
Db 1801 CGTAGACATGACCATTTATTAATGATGCTGGTGGGATGCTGTGTTCTAAATTAACCTAT 1860
QY 1861 TTTTCAAGGCTCTCTGTTATTTGAAGTATCATCTGGTGTGCTTAACTCTTTAAATTTG 1920
Db 1861 TTTTCAAGGCTCTCTGTTATTTGAAGTATCATCTGGTGTGCTTAACTCTTTAAATTTG 1920
QY 1921 TATATATTACTGTTAGTAAATTAATAATCAATATCCATATCCATATTAATTAAGTCA 1980
Db 1921 TATATATTACTGTTAGTAAATTAATAATCAATATCCATATCCATATTAATTAAGTCA 1980
QY 1981 ATATCTGCTCTTTGTATAGTATCATGAAATTCATAAAATTTATTTATCTCTTTATAGAA 2040
Db 1981 ATATCTGCTCTTTGTATAGTATCATGAAATTCATAAAATTTATTTATCTCTTTATAGAA 2040
QY 2041 TAAAGATTAAATATGTTTAAAAAAA 2065
Db 2041 TAAAGATTAAATATGTTTAAAAAAA 2065

RESULT 3
AX697151
LOCUS AX697151 2065 bp DNA linear PAT 02-APR-2003
DEFINITION Sequence 219 from Patent WO0078961.
ACCESSION AX697151
VERSION AX697151.1 GI:29498104
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Ferrara,N., Stewart,T.A., Williams,P.M., Baker,K.P., Desnoyers,L.,
Eaton,D.L., Gao,W.Q., Pan,J., Botstein,D., Fong,S., Goddard,A.,
Godowski,P.J., Gurney,A.L., Smith,V., Tumas,D., Wood,W.I.,
Grimaldi,C.J., Hillan,K.J., Paoni,N.F., Roy,M.A. and Watanabe,C.K.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0078961-A 219 28-DEC-2000;
Genentech Inc. (US)
FEATURES
source location/Qualifiers
1..2065
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 2065; DB 6; Length 2065;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1081 GTCTGAACCAATACAAATTTTCACAGATATTTCGTGTGCTCTTTTCAGTATATTTGGA 1140
Db 1081 GTCTGAACCAATACAAATTTTCACAGATATTTCGTGTGCTCTTTTCAGTATATTTGGA 1140
QY 1141 TTGGGACTCTAAGCAGATAATACCTATGCTTAATATGTAACAGTCAAAAGCTGTCTGCAAG 1200
Db 1141 TTGGGACTCTAAGCAGATAATACCTATGCTTAATATGTAACAGTCAAAAGCTGTCTGCAAG 1200
QY 1201 ACTTATTTCTGAATTTCTCTTCTGGGATTTACTGAAATAGTTTACAGATGTGGAATTTTATT 1260
Db 1201 ACTTATTTCTGAATTTCTCTTCTGGGATTTACTGAAATAGTTTACAGATGTGGAATTTTATT 1260
QY 1261 TGTCTAGTTTAAAGACTGCAACACAGCTCTAAGGATAGAAATCTCTAAAGTCTGAC 1320
Db 1261 TGTCTAGTTTAAAGACTGCAACACAGCTCTAAGGATAGAAATCTCTAAAGTCTGAC 1320
QY 1321 TTCAATCAACGGTGTAGTGTATCTGCAAGAACTGTATCTGTTTGGTCTCTGTAAGAACTTGGATTTT 1380
Db 1321 TTCAATCAACGGTGTAGTGTATCTGCAAGAACTGTATCTGTTTGGTCTCTGTAAGAACTTGGATTTT 1380
QY 1381 ATATTTCTTTTATTTCTTTTGGAAATAGTTTGTGTTTGGTCTCTGTAAGAACTTGGATTTT 1440
Db 1381 ATATTTCTTTTATTTCTTTTGGAAATAGTTTGTGTTTGGTCTCTGTAAGAACTTGGATTTT 1440
QY 1441 TTTTTCAGTAACCTGTAATTTATGTTTCTCTTAAATAAGTAATGATGCTTGGCCAC 1500
Db 1441 TTTTTCAGTAACCTGTAATTTATGTTTCTCTTAAATAAGTAATGATGCTTGGCCAC 1500
QY 1501 AAATTTACCTTGACTACGATATATCGACATGACTCTCTCAAAAAAAGAAAGTCTTCA 1560
Db 1501 AAATTTACCTTGACTACGATATATCGACATGACTCTCTCAAAAAAAGAAAGTCTTCA 1560
QY 1561 TAGTTGTATTTAATTTGTATGTAAGAGTCAATTTTCCAGTTTATATTTTCTAAGA 1620
Db 1561 TAGTTGTATTTAATTTGTATGTAAGAGTCAATTTTCCAGTTTATATTTTCTAAGA 1620
QY 1621 AGAAGAAATAGATCATATAATCTGCAAGAAAGTCTCTTACCCAAATCTAAGTGTCTCA 1680
Db 1621 AGAAGAAATAGATCATATAATCTGCAAGAAAGTCTCTTACCCAAATCTAAGTGTCTCA 1680
QY 1681 ATCCCTGAGCTCAGCAAAAACAGCTCCCTCCGAGGGAATCTTATCTTATTTATGCTCAA 1740
Db 1681 ATCCCTGAGCTCAGCAAAAACAGCTCCCTCCGAGGGAATCTTATCTTATTTATGCTCAA 1740
QY 1741 CTTTAAATTAATGATGATAATAACACATTTTAAATAACCTTAAGCTTTTCTTTC 1800
Db 1741 CTTTAAATTAATGATGATAATAACACATTTTAAATAACCTTAAGCTTTTCTTTC 1800
QY 1801 CGTAGACATGACCACTTTATTAATCTGCTGGTGGGATGCTGTTTCTTATTAACCTAT 1860
Db 1801 CGTAGACATGACCACTTTATTAATCTGCTGGTGGGATGCTGTTTCTTATTAACCTAT 1860
QY 1861 TTTTCAAGCTTCTGTTGTAATTTGAAGTATCATCTGTTTGTCTTAACTTAACTTAAATG 1920
Db 1861 TTTTCAAGCTTCTGTTGTAATTTGAAGTATCATCTGTTTGTCTTAACTTAACTTAAATG 1920
QY 1921 TATATATTTATCTGTTTGTAGTATAATTAATTAATTAATTAATTAATTAATTAATTAATG 1980
Db 1921 TATATATTTATCTGTTTGTAGTATAATTAATTAATTAATTAATTAATTAATTAATG 1980
QY 1981 ATATCTTGTCTTTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
Db 1981 ATATCTTGTCTTTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
QY 2041 TAAAGATTAATATATATGTTTAAAAA 2065
Db 2041 TAAAGATTAATATATGTTTAAAAA 2065
```

RESULT 4
AY358527 2065 bp mRNA linear PRI 03-OCT-2003
LOCUS
DEFINITION Homo sapiens clone DNA66526 GSGR718 (UNQ718) mRNA, complete cds.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY358527
AY358527.1 GI:371821176
FLI_CDNA:
Homo sapiens (human)

REFERENCE
AUTHORS

1 (bases 1 to 2065)
Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,
Chen,J., Chow,B., Chui,C., Crowley,C., Curreli,B., Deuel,B.,
Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,
Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S.,
Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C.,
Schoenfeld,J., Seehagiri,S., Simmons,L., Singh,J., Smith,V.,
Stinson,J., Vats,A., Vandlen,R., Watanabe,C., Weiland,D., Woods,K.,
Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z.,
Goddard,A., Wood,W.I. and Godowski,P.
The Secreted Protein Discovery Initiative (SPDI), a Large-Scale
Effort to Identify Novel Human Secreted and Transmembrane Proteins:
A Bioinformatics Assessment
Genome Res. 13 (10), 2265-2270 (2003)

TITLE

12975309
2. (bases 1 to 2065)
Clark,H.F.
Direct Submission
Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,
Inc., 1 DNA Way, South San Francisco, CA 94080, USA

JOURNAL
PUBMED
REFERENCE

12975309
2. (bases 1 to 2065)
Clark,H.F.
Direct Submission
Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,
Inc., 1 DNA Way, South San Francisco, CA 94080, USA

AUTHORS
JOURNAL

12975309
2. (bases 1 to 2065)
Clark,H.F.
Direct Submission
Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,
Inc., 1 DNA Way, South San Francisco, CA 94080, USA

FEATURES
source

1.2065
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DNA66526"
1.2065
/locus_tag="UNQ718"
337.942
/locus_tag="UNQ718"
/notes="PRO1382"
/codon_start=1
/product="GSGR718"
/protein_id="AAQ8891.1"
/db_xref="GI:371821177"
/translation="MGSRRRLSAVPAVLVLTLPGLVMAQNDTPIVLEKCLVVC
DSNPATDSKSSSSPLGISVRAANSKVFSAVSTNHPSEMSKTRIIYFDQILNVV
GNFTLESFVAPRKGIYSFPHVIVYOSQTOVNLMLNGKPVISAFAGKDVTEA
ANGVLLYLCKEDKIVLKLEKLNVLGVGWQYSIFSGLVFLP"

gene
CDS

1.2065
/locus_tag="UNQ718"
337.942
/locus_tag="UNQ718"
/notes="PRO1382"
/codon_start=1
/product="GSGR718"
/protein_id="AAQ8891.1"
/db_xref="GI:371821177"
/translation="MGSRRRLSAVPAVLVLTLPGLVMAQNDTPIVLEKCLVVC
DSNPATDSKSSSSPLGISVRAANSKVFSAVSTNHPSEMSKTRIIYFDQILNVV
GNFTLESFVAPRKGIYSFPHVIVYOSQTOVNLMLNGKPVISAFAGKDVTEA
ANGVLLYLCKEDKIVLKLEKLNVLGVGWQYSIFSGLVFLP"

ORIGIN

Query Match 100.0%; Score 2065; DB 9; Length 2065;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGAATGTAGGGTTTGTATGACTTTAGATGTTAGGAACCGAGTGGGTGAGGGGCC 60
Db 1 GTGAATGTAGGGTTTGTATGACTTTAGATGTTAGGAACCGAGTGGGTGAGGGGCC 60
QY 61 CAGGCGAGGCTGATTTCTTGGCGGAGGAGTAGGTAAAGGTTCTTCATGAGTCCCT 120
Db 61 CAGGCGAGGCTGATTTCTTGGCGGAGGAGTAGGTAAAGGTTCTTCATGAGTCCCT 120
QY 121 AAAGGACAAAGGTAACAGAGGAGAGCTCGAGGGGAGAGCTTTCGACTTCAAGCCA 180
Db 121 AAAGGACAAAGGTAACAGAGGAGAGCTCGAGGGGAGAGCTTTCGACTTCAAGCCA 180
QY 181 CAGAAATGTGAAGTGTGCGCGCGCCCGCGCGCTCTCTGACAGCGTGTGACCTA 240
Db 181 CAGAAATGTGAAGTGTGCGCGCGCCCGCGCGCTCTCTGACAGCGTGTGACCTA 240
QY 241 GCGCTAGCATCTTCCCGAGCACCGGGATCCCGGGGTAGGAGCGGCGGGGAGCACC 300
Db 241 GCGCTAGCATCTTCCCGAGCACCGGGATCCCGGGGTAGGAGCGGCGGGGAGCACC 300
QY 301 AGCGCCAGCGGCTGCGGCTGCCCAACACGCTCACCATGGGCTCCGGGCGCGGCGCTG 360

301 AGCGGAGCGGCTGGCGTGGCCACACGSCCTCACCATGGCTCGCGGCGCGGCGCTG 360
361 TCCGGGGTCCCGCGCGTGTCTGCTGCTCCTCAGCTGCGCGGGGCTGCGCTGCGGACAG 420
361 TCCGGGGTCCCGCGCGTGTCTGCTGCTCAGCTGCGCGGGGCTGCGCGGCTGCGGACAG 420
421 AACGACACGAGACCCATCGTGTGAGGACCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTG 480
421 AACGACACGAGACCCATCGTGTGAGGACCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTG 480
481 GCCAGGACTCCAAAGGCTCTCTCCCTCCCGCTGGGGATATCGGTGCGGCGGCGCCAC 540
481 GCCAGGACTCCAAAGGCTCTCTCTCCCTCCCGCTGGGGATATCGGTGCGGCGGCGCCAC 540
541 TCCAAAGTCCGCTTCTCGCGCGTGGGAGCACCAACACGAGCCATCCGAGATGAGCAAC 600
541 TCCAAAGTCCGCTTCTCGCGCGTGGGAGCACCAACACGAGCCATCCGAGATGAGCAAC 600
601 AAGAGCGCATCATTTTACTTCGATCAGATCTCTGCTGATGCTGGGTAATTTTTCACATTG 660
601 AAGAGCGCATCATTTTACTTCGATCAGATCTCTGCTGATGCTGGGTAATTTTTCACATTG 660
661 GAGTCTGCTTTGTAGCACCAAGAAAGAAATTTACAGTTTTCAGTTTTCAGTTTTCAGTTT 720
661 GAGTCTGCTTTGTAGCACCAAGAAAGAAATTTACAGTTTTCAGTTTTCAGTTTTCAGTTT 720
721 GTCTACGAGCCTATCCAGTTTAACTTGATTTTAAATGGAACCAAGTAAATCT 780
721 GTCTACGAGCCTATCCAGTTTAACTTGATTTTAAATGGAACCAAGTAAATCT 780
781 GCCTTTGGGGGACCAAGATGTTTACTGTAAGCTGCAAGATGCTGCTGCTGCTGCTGCTG 840
781 GCCTTTGGGGGACCAAGATGTTTACTGTAAGCTGCAAGATGCTGCTGCTGCTGCTGCTG 840
841 CTAGATAAAGAGGTAAGGTTTACCTAAATCTGAGAAAGTAAATTTGCTGGAGGCTGG 900
841 CTAGATAAAGAGGTAAGGTTTACCTAAATCTGAGAAAGTAAATTTGCTGGAGGCTGG 900
901 CAGTATTCACGTTTCTGCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
901 CAGTATTCACGTTTCTGCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
961 TGTTTCATCCAGTGGAGTACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
961 TGTTTCATCCAGTGGAGTACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
1021 TGGATTGATGCTTTTATTTGCTTCTCATGCTGCTGATGATGCTTAAAGGATCTAGCCT 1080
1021 TGGATTGATGCTTTTATTTGCTTCTCATGCTGCTGATGATGCTTAAAGGATCTAGCCT 1080
1081 GTCTGACCAATACAAATTTACAGATTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
1081 GTCTGACCAATACAAATTTACAGATTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
1141 TTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTAACAGTCAAAAGCTGCTGCAAG 1200
1141 TTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTAACAGTCAAAAGCTGCTGCAAG 1200
1201 ACTTATCTGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
1201 ACTTATCTGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
1261 TGTATTGTTTAAAGACTGCAACCGCTTAAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1320
1261 TGTATTGTTTAAAGACTGCAACCGCTTAAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1320
1321 TTCAATCAACGTTAGTGTGATCTGCAAAAGAACTGTAAGTGTGTAATATTAATTTGATT 1380
1321 TTCAATCAACGTTAGTGTGATCTGCAAAAGAACTGTAAGTGTGTAATATTAATTTGATT 1380
1381 ATATTGTTTATCTTTGGAATAGTTTGTGCTTCTGTAAGAACTTGGATTTT 1440

Db 1381 ATATTGTTTATTTATTCCTTTGGAATTAGTTTGTGTTGGTTCTTGTAAGAACTGGATTTT 1440
QY 1441 TTTTTCAGTAACCTGGTATTATGTTTCTTAAATAAGGTAATGAATGGCTTCCGCCAC 1500
Db 1441 TTTTTCAGTAACCTGGTATTATGTTTCTTAAATAAGGTAATGAATGGCTTCCGCCAC 1500
QY 1501 AAATTTACCTTACGATATCATCGATATCATCGATATCATCGATATCATCGATATCATCG 1560
Db 1501 AAATTTACCTTACGATATCATCGATATCATCGATATCATCGATATCATCGATATCATCG 1560
QY 1561 TAGTTGTTTAAATTTGATATGTAAGAGTCAATATTTTCCAAAGTATATTTTCTAAGA 1620
Db 1561 TAGTTGTTTAAATTTGATATGTAAGAGTCAATATTTTCCAAAGTATATTTTCTAAGA 1620
QY 1621 AGAAGATAGATCAATAATCTGACAGAAAGTGTCTTACCCAAATCTAAGTGCTCA 1680
Db 1621 AGAAGATAGATCAATAATCTGACAGAAAGTGTCTTACCCAAATCTAAGTGCTCA 1680
QY 1681 ATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGGAATCTTATCTTTTATTTGCTCAA 1740
Db 1681 ATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGGAATCTTATCTTTTATTTGCTCAA 1740
QY 1741 CTTTAAATTAAGTATGATAATAACCACTTTTATTAAGGTTTCTTTTCTTTTCTTCT 1800
Db 1741 CTTTAAATTAAGTATGATAATAACCACTTTTATTAAGGTTTCTTTTCTTTTCTTCT 1800
QY 1801 CGTAGACATGACCACTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1801 CGTAGACATGACCACTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
QY 1861 TTTTCAAGGCTTCTGTTGTTTGAAGTATCATCTGCTGTTTGGCTTAACTTCTTTAAATTTG 1920
Db 1861 TTTTCAAGGCTTCTGTTGTTTGAAGTATCATCTGCTGTTTGGCTTAACTTCTTTAAATTTG 1920
QY 1921 TATATTTATCTGTTTGTAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980
Db 1921 TATATTTATCTGTTTGTAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980
QY 1981 ATATCTTCTTCTTTGATAGTCTATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
Db 1981 ATATCTTCTTCTTTGATAGTCTATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
QY 2041 TAAAGATTAATATATGTTTAAAAAAA 2065
Db 2041 TAAAGATTAATATATGTTTAAAAAAA 2065

RESULT 5
HSJ885A10/c

LOCUS Human DNA sequence from clone RP5-885A10 on chromosome 20 Contains
DEFINITION a gene for a protein similar to cerebellin precursor, ESTs, STSS,
GSs, and a CpG island, complete sequence.
ACCESSION AL117383
VERSION AL117383.19 GI:7406730
KEYWORDS HTG; cerebellin; CpG Island.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 117854)
AUTHORS Laird, G.
TITLE Direct Submission
JOURNAL Submitted (26-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT requests: clonequery@sanger.ac.uk
On Apr 3, 2000 this sequence version replaced gi:7264032.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-885A10 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone RP5-885A10. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP5-885A10 is at 1 in this sequence. The true left end of clone RP11-380D15 is at 11755 in this sequence. The true right end of clone RP11-458E15 is at 27524 in this sequence.

FEATURES

source	Location/Qualifiers
	1..117854
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="20"
	/clones="RP5-885A10"
	/clone_lib="RPCI-5"
	585..740
repeat_region	/note="78 copies 2 mer aa 59% conserved"
repeat_region	594..649
repeat_region	/note="7 copies 8 mer aagaaag 76% conserved"
repeat_region	600..643
repeat_region	/note="11 copies 4 mer gga 79% conserved"
misc_feature	761..1245
gene	/note="match: STS: Em:AA869507"
gene	complement(1551..7478)
mRNA	/gene="dJ885A10.1"
	/product="dJ885A10.1 (similar to cerebellin precursor)"
	/gene="dJ885A10.1"
	/note="match: cDNAs: Em:M58583
	match: ESTs: Em:AV72278
	/evidence=not_experimental
CDS	complement(join(1864..2061,4038..4154,7188..7478))
	/gene="dJ885A10.1"
	/note="match: proteins: Sw:P98087 Sw:P23435"
	/codon_start=1
	/evidence=not_experimental
	/product="dJ885A10.1 (similar to cerebellin precursor)"
	/protein_id="CAB90173.1"
	/db_xref="GI:7708438"
	/db_xref="SWISS-PROT:Q9NTU7"
	/translation="MSGGRALSAVPAVLVLPLGLPWAQNDTPVLGKCLVVC DSNPATDKSSSPSPGIGSVRAANSKVAFAVSRTHBPSMKNKTRITFYDQILVNV GNFTLESVFAVRKGYISFHFVIKYQSTIQVNLMLNGKPVISAFAGDKDVTREA ATNGVLVLDKEDKVLKLEKGLVGVGQWSTPSGLVFLP"
repeat_region	2642..2958
	/note="AluJo repeat: matches 1..306 of consensus"
misc_feature	complement(2788..3347)
	/gene="dJ885A10.1"
misc_feature	/note="match: GSS: Em:AQ530830"
	complement(2972..3470)
	/gene="dJ885A10.1"
misc_feature	/note="match: GSS: Em:AQ317752"
	complement(3040..3495)
	/gene="dJ885A10.1"
	/note="match: GSS: Em:AQ408825"
	3341..3840
misc_feature	/note="match: GSS: Em:AQ514309"
misc_feature	3499..4061
misc_feature	/note="match: STS: Em:HS267E11T"
	3608..3929
	/note="match: GSS: Em:AQ594830"
	4245..4280
repeat_region	/note="9 copies 4 mer ctcc 91% conserved"
repeat_region	4282..4325
misc_feature	/note="11 copies 4 mer tctt 81% conserved"
	complement(6245..6743)
	/gene="dJ885A10.1"
	/note="match: GSS: Em:B70819"
	6729..7220
	/note="match: GSS: Em:AQ676409"
	6855..8929
	/note="CpG island"
	/evidence=not_experimental
repeat_region	9008..9049
	/note="7 copies 6 mer aaaaat 81% conserved"
repeat_region	9009..9052
repeat_region	/note="22 copies 2 mer aa 79% conserved"
	9770..9892
	/note="FLAM A repeat: matches 12..108 of consensus"
	10273..10302
	/note="15 copies 2 mer aa 96% conserved"
	12095..12694
	/note="LIP4 repeat: matches 5538..6140 of consensus"
	13449..13623
	/note="5 copies 35 mer 77% conserved"
	13484..13599
	/note="29 copies 4 mer tctc 75% conserved"
	14187..14481
	/note="AluY repeat: matches 1..294 of consensus"
	15140..15329
	/note="MIR repeat: matches 62..262 of consensus"
	16419..16484
	/note="MIR repeat: matches 124..192 of consensus"
	16438..17076
	/note="match: GSS: Em:AQ391974"
	16830..16899
	/note="2 copies 35 mer 92% conserved"
	17052..17111
	/note="L2 repeat: matches 2651..2710 of consensus"
	17910..17987
	/note="39 copies 2 mer tt 64% conserved"
	19009..19305
	/note="AluX repeat: matches 1..298 of consensus"
	19337..19686
	/note="THE1B repeat: matches 1..364 of consensus"
	20347..20640
	/note="AluX repeat: matches 1..294 of consensus"
	20747..20874
	/note="MER5B repeat: matches 1..145 of consensus"
	22332..22267
	/note="6 copies 6 mer cacaca 86% conserved"
	22233..22264
	/note="8 copies 4 mer acac 90% conserved"
	complement(22555..23071)
misc_feature	/note="match: GSS: Em:AQ451152"
misc_feature	complement(22573..23081)
	/note="match: GSS: Em:AQ053765"
	22757..23455
	/note="L2 repeat: matches 1992..2710 of consensus"
repeat_region	24097..24357
	/note="AluSq repeat: matches 1..261 of consensus"
misc_feature	complement(24374..24608)
misc_feature	/note="match: GSS: Em:AQ677215"
	24394..24844
	/note="match: GSS: Em:AQ762377"
	24839..25272

```

misc_feature      /note="match: GSS: Em:B58489"
24845..25393
misc_feature      /note="match: GSS: Em:AQ899089"
24942..25279
repeat_region    /note="match: GSS: Em:AQ222882"
25892..26161
repeat_region    /note="L2 repeat: matches 2090..2384 of consensus"
26231..26402
repeat_region    /note="MER3 repeat: matches 5..176 of consensus"
26407..27166
misc_feature      /note="LIPAL3 repeat: matches 5388..6155 of consensus"
complement(27157..27515)
repeat_region    /note="match: GSS: Em:AQ582450"
27175..27459
repeat_region    /note="AluSx repeat: matches 1..285 of consensus"
27585..27738
repeat_region    /note="AluY repeat: matches 147..300 of consensus"
27744..27934
repeat_region    /note="L2 repeat: matches 2555..2750 of consensus"
28859..29029
repeat_region    /note="MIR repeat: matches 1..218 of consensus"
complement(29154..29702)
misc_feature      /note="match: GSS: Em:AQ385339"
29329..29553
repeat_region    /note="MIR repeat: matches 32..262 of consensus"
29593..30173
misc_feature      /note="match: GSS: Em:AQ052168"
match: SFS: Em:AQ052168"
29658..29737
repeat_region    /note="MER96 repeat: matches 1..71 of consensus"

```

```

Query Match      63.8%; Score 1318.4; DB 9; Length 117854;
Best Local Similarity 99.9%; Pred.No. 28-222;
Matches 1319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 740 TCAGGTTAACTGATGTTAAATGGAACACCAAGTATATCTGCCCTTGGGGGACAAAG 799
DB 2066 TCTAGGTTAACTGATGTTAAATGGAACACCAAGTATATCTGCCCTTGGGGGACAAAG 2007
QY 800 ATGTTACTCGTGAAGTCCAGCAATGGTGTCTCTCTACCTAGATTAAGAGGATAAGG 859
DB 206 ATGTTACTCGTGAAGTCCAGCAATGGTGTCTCTCTACCTAGATTAAGAGGATAAGG 1947
QY 860 TTACTCTAACTGAGAAAGTAAATTTGGTGGAGGCTGGCAGTATCCACGTTTCG 919
DB 1946 TTACTCTAACTGAGAAAGTAAATTTGGTGGAGGCTGGCAGTATCCACGTTTCG 1887
QY 920 GCTTCTGGTGTCCCTATAGGATTCAAATTTCTCCATGATGTTTCATCCAGGTGAGGA 979
DB 1886 GCTTCTGGTGTCCCTATAGGATTCAAATTTCTCCATGATGTTTCATCCAGGTGAGGA 1827
QY 980 TGACCCACTCTGAGTATTGGAAGATCATTTTTCATCANTGGATTGATGCTTTTATT 1039
DB 1826 TGACCCACTCTGAGTATTGGAAGATCATTTTTCATCANTGGATTGATGCTTTTATT 1767
QY 1040 GGTTCATCGGTGGATATGATTTCTAGGATTCAGGATTCAGCTGTCTGACCAATCAAAAT 1099
DB 1766 GGTTCATCGGTGGATATGATTTCTAGGATTCAGGATTCAGCTGTCTGACCAATCAAAAT 1707
QY 1100 TTCACAGTATTATGCTGTGTCTGTTTCAGTATATTTGGATTGGGACTCTAAGCAGATA 1159
DB 1706 TTCACAGTATTATGCTGTGTCTGTTTCAGTATATTTGGATTGGGACTCTAAGCAGATA 1647
QY 1160 ATACCTATGCTAAATGTAACAGTCAAAAGCTGTCTGCAAGCTTATTCGAAATTCATT 1219
DB 1646 ATACCTATGCTAAATGTAACAGTCAAAAGCTGTCTGCAAGCTTATTCGAAATTCATT 1587
QY 1220 TCCTGGGATTACTGAATTTAGTATACAGATGCGAATTTTATTTGTTTATGTTTAAAGACT 1279
DB 1586 TCCTGGGATTACTGAATTTAGTATACAGATGCGAATTTTATTTGTTTATGTTTAAAGACT 1527
QY 1280 GGCACACCGGTCTAAGGATTAGAAACCTTAAAGTCTGACCTCAATCAACGCTTAGTGT 1339

```

```

DB 1526 GGCACACCGGTCTAAGGATTAGAAACCTTAAGTTCGACTTCAATCAACGGTTAGTGT 1467
QY 1340 GATACCTGCCAAGAACTGTATACCTGTGTGTAAATATATTAATTAATTTGTTTATTCCTT 1399
DB 1466 GATACCTGCCAAGAACTGTATACCTGTGTGTAAATATATTAATTAATTTGTTTATTCCTT 1407
QY 1400 TCGAATTAAGTGTGTTTGGTTCTGTGTAATAAACTCTGATTTTTCAGTAACTGGTAT 1459
DB 1406 TCGAATTAAGTGTGTTTGGTTCTGTGTAATAAACTCTGATTTTTCAGTAACTGGTAT 1347
QY 1460 TATGTTTCTCTTAAAAATAAGTAATGAATGCTTCCACAAATTTACCTTTGACTACGA 1519
DB 1346 TATGTTTCTCTTAAAAATAAGTAATGAATGCTTCCACAAATTTACCTTTGACTACGA 1287
QY 1520 TATCATCCACATGACTCTCTCAAAAAAAGAAAGCTTCATAGTGTATTTTAAATTTGA 1579
DB 1286 TATCATCCACATGACTCTCTCAAAAAAAGAAAGCTTCATAGTGTATTTTAAATTTGA 1227
QY 1580 TATGTGAAGAGTCAATATTTTCCAAAGTTATATTTCTAAGAAGAAGATAGATCAATAAT 1639
DB 1226 TATGTGAAGAGTCAATATTTTCCAAAGTTATATTTCTAAGAAGAAGATAGATCAATAAT 1167
QY 1640 CTGACAAGGAAAGTGTGCTTACCCAAATCTAAGTGTCTCAATCCCTGAGCCTCAGCAA 1699
DB 1166 CTGACAAGGAAAGTGTGCTTACCCAAATCTAAGTGTCTCAATCCCTGAGCCTCAGCAA 1107
QY 1700 ACAGCTCCCTCCGAGGGAATCTTATACCTTTATGCTCAACTTTAAATTAAGTATGA 1759
DB 1106 ACAGCTCCCTCCGAGGGAATCTTATACCTTTATGCTCAACTTTAAATTAAGTATGA 1047
QY 1760 TAAATACCACTTTTATTAATAAACCCTAAGGTTTTTTTTTCCGTAGACATGACCACTTTA 1819
DB 1046 TAAATACCACTTTTATTAATAAACCCTAAGGTTTTTTTTTCCGTAGACATGACCACTTTA 987
QY 1820 TTAACCTGTGGTGGAGTGTGCTTCTTAATTAATACCTATTTTCAAGGCTCTCTGTTGT 1879
DB 986 TTAACCTGTGGTGGAGTGTGCTTCTTAATTAATACCTATTTTCAAGGCTCTCTGTTGT 927
QY 1880 ATTTGAAGTATCATCTGTTTGGCTTAACTTTTAAATGCTATATATTTATCTGTTTAG 1939
DB 926 ATTTGAAGTATCATCTGTTTGGCTTAACTTTTAAATGCTATATATTTATCTGTTTAG 867
QY 1940 CTAATATTAATCAATATCCCATATCTAAATTTAGTGAATATCTTGTCTTTTGTATA 1999
DB 866 CTAATATTAATCAATATCCCATATCTAAATTTAGTGAATATCTTGTCTTTTGTATA 807
QY 2000 GGTCTATGATTCATAAAATTAATTTATGCTGTGTATAGATAAAGATTAATATATGTTA 2059
DB 806 GGTCTATGATTCATAAAATTAATTTATGCTGTGTATAGATAAAGATTAATATATGTTA 747

```

```

RESULT 6
BC047076
LOCUS Homo sapiens, clone IMAGE:5314390, mRNA.
DEFINITION BC047076
ACCESSION BC047076
VERSION BC047076.1 GI:28436739
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 926)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

```


cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Loulseged, H., Kowis, C. R., Sneed, A. J., Martin, R. G., Muzny, D. M., Nanavati, A. N., Gibbs, R. A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: TRAK Map: 91 Row: o Column: 23.

FEATURES

Location/Qualifiers
 1..926
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5314390"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH_MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

ORIGIN

Query Match 42.3%; Score 873; DB 9; Length 926;
 Best Local Similarity 99.8%; Pred. No. 4.3e-144;
 Matches 895; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1169 CTTAAATGTAACAGTCAAAAGCTGCTGCAAGACTTATTCGAATTCATTTCTCTGGGAT 1228
 DB 1 CTTAAATGTAACAGTCAAAAGCTGCTGCAAGACTTATTCGAATTCATTTCTCTGGGAT 60

QY 1229 TACTGAATAGTACAGATGGGAATTTATTTGTTTATTTTAAAGACTGGCAACGAG 1288
 DB 61 TACTGAATAGTACAGATGGGAATTTATTTGTTTATTTTAAAGACTGGCAACGAG 120

QY 1289 GCTTAAGGATTAGAAAACCTTAAGCTTCGACTTCAATCAACGGTGTAGTGATCTGCC 1348
 DB 121 GCTTAAGGATTAGAAAACCTTAAGCTTCGACTTCAATCAACGGTGTAGTGATCTGCC 180

QY 1349 AAAGAACTGTACTGTGTGTTAATATATGATTATATTTGTTTATTTCTTTGGAAATTAG 1408
 DB 181 AAAGAACTGTACTGTGTGTT-ATAATGATTATATTTGTTTATTTCTTTGGAAATTAG 239

QY 1409 TTTGTTTGGTCTCTGTAATAAAACCTTGGATTTTTTTTTCAGTAACCTGGTATTATGTTTC 1468
 DB 240 TTTGTTTGGTCTCTGTAATAAAACCTTGGATTTTTTTTTCAGTAACCTGGTATTATGTTTC 299

QY 1469 TCTTAAATTAAGTGAATGAATGCTTGGCCCAAAATTTACCTGTAGTACGATCATCGA 1528
 DB 300 TCTTAAATTAAGTGAATGAATGCTTGGCCCAAAATTTACCTGTAGTACGATCATCGA 359

QY 1529 CATGACTTCTCTCAAAAAAAGAAAGTCTTCATAGTGTATTTAATTTGATATGTGAAA 1588
 DB 360 CATGACTTCTCTCAAAAAAAGAAAGTCTTCATAGTGTATTTAATTTGATATGTGAAA 419

QY 1589 GAGTCATATTTTCCAAAGTTATATTTCTAAGAAAGAAATAGATCATATAATCTGCAAGG 1648
 DB 420 GAGTCATATTTTCCAAAGTTATATTTCTAAGAAAGAAATAGATCATATAATCTGCAAGG 479

QY 1649 AAAAAAGTCTCTACCCAAATCTAAGTGTCTCAATCCCTGAGCTCAGCAAAACAGCTCCC 1708
 DB 480 AAAAAAGTCTCTACCCAAATCTAAGTGTCTCAATCCCTGAGCTCAGCAAAACAGCTCCC 539

QY 1709 CTCGAGGGAATCTTATATCTTTATTTGCTCAACTTTAATTAATGATTGATAATAACCA 1768
 DB 540 CTCGAGGGAATCTTATATCTTTATTTGCTCAACTTTAATTAATGATTGATAATAACCA 599

QY 1769 CTTTATTAACCTTAAGGTTTTTTTTTTTTCGTTAGACATGACCACTTTATTAACGTGT 1828
 |||||||

Db 600 CTTTATTAAAAACCTAAGG-TTTTTTTTTTTTCCGTAGACATGACCACTTTTATTAACTGGT 658
 QY 1829 GGTGGGATGCTGTGTTTCTTAATATATACCTATTTTTCAGAGCTCTCTGTATTGAGT 1888
 Db 659 GGTGGGATGCTGTGTTTCTTAATATATACCTATTTTTCAGAGCTCTCTGTATTGAGT 718
 QY 1889 ATCATCTGGTTTTTGCCTTAACCTCTTTAAATGTATATATTTATCTGTTTACCTAATATTA 1948
 Db 719 ATCATCTGGTTTTTGCCTTAACCTCTTTAAATGTATATATTTATCTGTTTACCTAATATTA 778
 QY 1949 AATTCAAAATATCCCATATCTAAATTTAGTGAATATCTGTCTTTTGTATAGGTCATATG 2008
 Db 779 AATTCAAAATATCCCATATCTAAATTTAGTGAATATCTGTCTTTTGTATAGGTCATATG 838
 QY 2009 AATTCAAAATATTTATGTCCTGTATAGATAAAGATTAATATATATGTTTAAAAAAA 2065
 Db 839 AATTCAAAATATTTATGTCCTGTATAGATAAAGATTAATATATGTTTAAAAAAA 895

RESULT 7

AC015860/c
 LOCUS AC015860 248469 bp DNA linear HTG 13-JUL-2000
 DEFINITION Homo sapiens clone RP11-42A1, LOW-PASS SEQUENCE SAMPLING.
 AC015860
 AC015860.3 GI:9157673
 VERSION HTG; HTGS PHASE0.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 248469)
 Homo sapiens, clone RP11-42A1
 Unpublished
 2 (bases 1 to 248469)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
 Baldwin, J., Barna, N., Becker, R., Boguslavsky, L., Bouckghalter, B.,
 Brown, A., Castle, A., Collangelo, M., Collins, S., Collymore, A.,
 Cooke, P., DeRellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
 Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gage, D.,
 Galagan, J., Gardyna, S., Grant, G., Hegos, B., Heaford, A., Horton, L.,
 Howland, J., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,
 Lehocsky, J., Lie, C., Locke, K., Macdonald, P., Marquis, N.,
 McSwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meidrin, J.,
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testay, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

COMMENT

JOURNAL

AUTHORS

REFERENCE

* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 863: contig of 863 bp in length
864 963: gap of 100 bp
964 1841: contig of 878 bp in length
1842 1941: gap of 100 bp
1942 2831: contig of 890 bp in length
2832 2931: gap of 100 bp
2932 3755: contig of 824 bp in length
3756 3855: gap of 100 bp
3856 4750: contig of 895 bp in length
4751 4850: gap of 100 bp
4851 5744: contig of 894 bp in length
5745 5844: gap of 100 bp
5845 6713: contig of 869 bp in length
6714 6813: gap of 100 bp
6814 7715: contig of 902 bp in length
7716 7815: gap of 100 bp
7817 8692: contig of 877 bp in length
8693 8792: gap of 100 bp
8793 9674: contig of 882 bp in length
9675 9774: gap of 100 bp
9776 10656: contig of 882 bp in length
10657 10756: gap of 100 bp
10757 11630: contig of 874 bp in length
11631 11730: gap of 100 bp
11731 12610: contig of 880 bp in length
12611 12710: gap of 100 bp
12712 13622: contig of 912 bp in length
13623 13722: gap of 100 bp
13723 14592: contig of 870 bp in length
14593 14692: gap of 100 bp
14693 15567: contig of 875 bp in length
15568 15667: gap of 100 bp
15669 16420: contig of 753 bp in length
16421 16520: gap of 100 bp
16521 17444: contig of 924 bp in length
17445 17544: gap of 100 bp
17546 18422: contig of 878 bp in length
18423 18522: gap of 100 bp
18523 19463: contig of 941 bp in length
19464 19563: gap of 100 bp
19564 20472: contig of 909 bp in length
20473 20572: gap of 100 bp
20573 21444: contig of 872 bp in length
21445 21544: gap of 100 bp
21546 22434: contig of 890 bp in length
22435 22534: gap of 100 bp
22535 23414: contig of 880 bp in length
23415 24356: contig of 842 bp in length
24357 24456: gap of 100 bp
24457 25329: contig of 873 bp in length
25330 25429: gap of 100 bp
25430 26315: contig of 886 bp in length
26316 26415: gap of 100 bp
26416 27322: contig of 907 bp in length
27323 27422: gap of 100 bp
27423 28291: contig of 869 bp in length
28292 28391: gap of 100 bp
28392 29255: contig of 864 bp in length
29256 29355: gap of 100 bp
29356 30234: contig of 879 bp in length
30235 30334: gap of 100 bp
30335 31224: contig of 890 bp in length
31225 31324: gap of 100 bp
31325 32255: contig of 931 bp in length
32256 32355: gap of 100 bp
32356 33282: contig of 927 bp in length
33283 33382: gap of 100 bp
33383 34259: contig of 877 bp in length
34260 34359: gap of 100 bp

34360 35252: contig of 893 bp in length
35253 35352: gap of 100 bp
35353 36214: contig of 862 bp in length
36215 36314: gap of 100 bp
36315 37222: contig of 908 bp in length
37223 38320: gap of 100 bp
38321 38420: contig of 998 bp in length
38421 39306: contig of 886 bp in length
39307 39406: gap of 100 bp
39407 40276: contig of 870 bp in length
40277 40376: gap of 100 bp
40377 41271: contig of 895 bp in length
41272 41371: gap of 100 bp
41372 42255: contig of 884 bp in length
42256 42355: gap of 100 bp
42356 43225: contig of 870 bp in length
43226 43325: gap of 100 bp
43326 44186: contig of 861 bp in length
44187 44286: gap of 100 bp
44287 45182: contig of 896 bp in length
45183 45282: gap of 100 bp
45283 46191: contig of 909 bp in length
46192 46291: gap of 100 bp
46292 47171: contig of 880 bp in length
47172 47271: gap of 100 bp
47272 48137: contig of 866 bp in length
48138 48237: gap of 100 bp
48239 49176: contig of 939 bp in length
49177 49276: gap of 100 bp
49277 50214: contig of 938 bp in length
50215 50314: gap of 100 bp
50315 51073: contig of 759 bp in length
51074 51173: gap of 100 bp
51174 52022: contig of 849 bp in length
52023 52122: gap of 100 bp
52123 53012: contig of 890 bp in length
53013 53112: gap of 100 bp
53113 54013: contig of 901 bp in length
54014 54113: gap of 100 bp
54115 55000: contig of 887 bp in length
55001 55100: gap of 100 bp
55101 55972: contig of 872 bp in length
55973 56072: gap of 100 bp
56073 56948: contig of 876 bp in length
56949 57048: gap of 100 bp
57049 57945: contig of 897 bp in length
57946 58045: gap of 100 bp
58046 58929: contig of 884 bp in length
58930 59029: gap of 100 bp
59030 59927: contig of 898 bp in length
59928 60027: gap of 100 bp
60028 60922: contig of 895 bp in length
60923 61022: gap of 100 bp
61023 61918: contig of 896 bp in length
61919 62018: gap of 100 bp
62019 62930: contig of 912 bp in length
62931 63030: gap of 100 bp
63031 63978: contig of 948 bp in length
63979 64078: gap of 100 bp
64079 64999: contig of 921 bp in length
65000 65099: gap of 100 bp
65100 66010: contig of 911 bp in length
66011 66110: gap of 100 bp
66111 67066: contig of 956 bp in length
67067 67166: gap of 100 bp
67167 68147: contig of 981 bp in length
68148 68247: gap of 100 bp
68248 69185: contig of 938 bp in length
69186 69285: gap of 100 bp
69286 70218: contig of 933 bp in length
70219 70318: gap of 100 bp
70319 71216: contig of 898 bp in length

```

*      71217 71316: gap of 100 bp
*      71317 72243: contig of 927 bp in length

Query Match      36.7%; Score 758.8; DB 2; Length 248469;
Best local Similarity 84.1%; Pred. No. 1.6e-123;
Matches 786; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 984 CCACCTCGAGTTATTGGAGATCATTTTTCATCATTGGATTCATGCTTTTATTTGGTT 1043
    |||
Db 169077 CCACCTCGAGTTATTGGAGATCATTTTTCATCATTGGATTCATGCTTTTATTTGGTT 169018

QY 1044 TCTCATGGTGGATATGATTTCTAAGGATTCAGCCTGTCGAACCAATACAAAATTTC 1103
    |||
Db 169017 TCTCATGGTGGATATGATTTCTAAGGATTCAGCCTGTCGAACCAATACAAAATTTC 168958

QY 1104 CAGATTATTGTGTGCTGTTTCAGTATATTTTGGATTGGGATCTTAAGCAGATATATC 1163
    |||
Db 168957 CAGATTATTGTGTGCTGTTTCAGTATATTTTGGATTGGGATCTTAAGCAGATATATC 168898

QY 1164 CTATGCTTAAATGTAACAGTCAAAAGCTGTCGAAGACTTATTTCTGAAATTCATTTCCCT 1223
    |||
Db 168997 CTATGCTTAAATGTAACAGTCAAAAGCTGTCGAAGACTTATTTCTGAAATTCATTTCCCT 168838

QY 1224 GGGATTACTGAATTAAGTACAGATGTCGAATTTTATTTTGTAGTTTAAAGACTGGCA 1283
    |||
Db 168837 GGGATTACTGAATTAAGTACAGATGTCGAATTTTATTTTGTAGTTTAAAGACTGGCA 168778

QY 1284 ACAGGCTTAAGGATTAAGAACTCTAAAGTCTGACTCAATCAACGGTTAGTGATATC 1343
    |||
Db 168777 ACCAGGCTTAAGGATTAAGAACTCTAAAGTCTGACTCAATCAACGGTTAGTGATATC 168718

QY 1344 CTGCCAAGACTCTATACCTGTTTAAATATATATGATATATTTCTTTTATTCCTTTGGA 1403
    |||
Db 168717 CTGCCAAGACTCTATACCTGTTTAAATATATATGATATATTTCTTTTATTCCTTTGGA 168658

QY 1404 ATTAGTTTGTGTTCTTGTGTAAGAACTTGGATTTTTTTTTTTCAGTAACTGGTATTATG 1463
    |||
Db 168657 ATTAGTTTGTGTTCTTGTGTAAGAACTTGGATTTTTTTTTTTCAGTAACTGGTATTATG 168598

QY 1464 TTTTCTCTTAAATTAAGTAATGAATGCTTGCCCAAAATTTACCTTGACTAGATATC 1523
    |||
Db 168597 TTTTCTCTTAAATTAAGTAATGAATGCTTGCCCAAAATTTACCTTGACTAGATATC 168538

QY 1524 ATCCACATGACTCTCTCABAAAAAGAAAGCTTCATAGTTGATTTTATTTATTTATG 1583
    |||
Db 168537 ATCCACATGACTCTCTCABAAAAAGAAAGCTTCATAGTTGATTTTATTTATTTATG 168478

QY 1584 TGAAGAGTCTATTTTCCAAAGTTATATTTTCTAAGAGAAGAAATAGATCAATAAATCTGA 1643
    |||
Db 168477 TGAAGAGTCTATTTTCCAAAGTTATATTTTCTAAGAGAAGAAATAGATCAATAAATCTGA 168418

QY 1644 CAAGGAAAAAGTTGTTACCAAAATCTAAGTGTCTCAATCCTGAGCCTCAGCAAAACAG 1703
    |||
Db 168417 CAAGGAAAAAGTTGTTACCAAAATCTAAGTGTCTCAATCCTGAGCCTCAGCAAAACAG 168358

QY 1704 CTCCTCCCGAGGAAATCTTTATCTTTTATTTGCTCAACTTTTAAATTAATGATTAAT 1763
    |||
Db 168357 CTCCTCCCGAGGAAATCTTTATCTTTTATTTGCTCAAAATTTTAAATTTTAAATTTT 168298

QY 1764 AACCACTTTATTAATAAACCTAAGGTTTTTTTTTTTTCGCTAGACATGACCACTTTATTA 1823
    |||
Db 168297 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 168238

QY 1824 CTGGTGGGGAATGCTGTGTTTCTAATATATACCTATTTTTCAGGCTTCGTGTATTT 1883
    |||
Db 168237 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 168178

QY 1884 GAAGTATCATCTGGTTTTCCTTAACTCTTTAAAT 1918
    |||
Db 168177 AAGTTGACTTTGAGGGTCTCTTTGTGATGTTAAAT 168143

```

RESULT 8

AC012335
LOCUS
DEFINITION

Homo sapiens clone RP11-28A1, *** SEQUENCING IN PROGRESS ***, 58
unordered pieces.

AC012335
ACCESSION
VERSION
KEYWORDS
SOURCE

HTG; HTGS PHASE1

Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 183553)

AUTHORS

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE

2 (bases 1 to 183553)

JOURNAL

REFERENCE

AUTHORS

Unpublished
Direct Submission
Submitted (23-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 24, 2000 this sequence version replaced gi:6102665.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE

JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Center project name: L600

Center clone name: 28_A_1

Project Information

Center project name: L600

Center clone name: 28_A_1

NOTE: This is a 'working draft' sequence. It currently

consists of 58 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1318: contig of 1318 bp in length

1319: contig of 100 bp

1320: contig of 100 bp

1321: contig of 100 bp

1322: contig of 100 bp

1323: contig of 100 bp

1324: contig of 100 bp

1325: contig of 100 bp

1326: contig of 100 bp

1327: contig of 100 bp

1328: contig of 100 bp

1329: contig of 100 bp

1330: contig of 100 bp

1331: contig of 100 bp

1332: contig of 100 bp

1333: contig of 100 bp

1334: contig of 100 bp

1335: contig of 100 bp

1336: contig of 100 bp

1337: contig of 100 bp

1338: contig of 100 bp

1339: contig of 100 bp

1340: contig of 100 bp

1341: contig of 100 bp

1342: contig of 100 bp

1343: contig of 100 bp

```
* 13830 13929: gap of 100 bp
* 13930 15248: contig of 1319 bp in length
* 15249 15348: gap of 100 bp
* 15349 16757: contig of 1409 bp in length
* 16758 16858: gap of 100 bp
* 16859 18403: contig of 1546 bp in length
* 18404 18503: gap of 100 bp
* 18504 19522: contig of 1019 bp in length
* 19523 19622: gap of 100 bp
* 19623 21123: contig of 1501 bp in length
* 21124 21223: gap of 100 bp
* 21224 22242: contig of 1019 bp in length
* 22243 22342: gap of 100 bp
* 22343 23634: contig of 1292 bp in length
* 23635 23734: gap of 100 bp
* 23735 25389: contig of 1655 bp in length
* 25390 25489: gap of 100 bp
* 25490 26698: contig of 1208 bp in length
* 26699 26797: gap of 100 bp
* 26798 28096: contig of 1299 bp in length
* 28097 28196: gap of 100 bp
* 28197 29328: contig of 1132 bp in length
* 29329 29428: gap of 100 bp
* 29429 30679: contig of 1251 bp in length
* 30680 30779: gap of 100 bp
* 30780 32106: contig of 1327 bp in length
* 32107 32206: gap of 100 bp
* 32207 33847: contig of 1641 bp in length
* 33848 33947: gap of 100 bp
* 33948 35798: contig of 1851 bp in length
* 35799 35898: gap of 100 bp
* 35899 37908: contig of 2010 bp in length
* 37909 38008: gap of 100 bp
* 38009 39061: contig of 1053 bp in length
* 39062 39161: gap of 100 bp
* 39162 40468: contig of 1307 bp in length
* 40469 40568: gap of 100 bp
* 40569 43284: contig of 2716 bp in length
* 43285 43384: gap of 100 bp
* 43385 44719: contig of 1335 bp in length
* 44720 44819: gap of 100 bp
* 44820 46467: contig of 1648 bp in length
* 46468 46567: gap of 100 bp
* 46568 48871: contig of 2304 bp in length
* 48872 48971: gap of 100 bp
* 48972 50204: contig of 1233 bp in length
* 50205 50304: gap of 100 bp
* 50305 52244: contig of 1940 bp in length
* 52245 52344: gap of 100 bp
* 52345 53401: contig of 1057 bp in length
* 53402 53501: gap of 100 bp
* 53502 55091: contig of 1590 bp in length
* 55092 55191: gap of 100 bp
* 55192 58084: contig of 2893 bp in length
* 58085 58184: gap of 100 bp
* 58185 60261: contig of 2077 bp in length
* 60262 60361: gap of 100 bp
* 60362 61645: contig of 1284 bp in length
* 61646 61745: gap of 100 bp
* 61746 63437: contig of 1692 bp in length
* 63438 63537: gap of 100 bp
* 63538 64887: contig of 1250 bp in length
* 64888 64987: gap of 100 bp
* 64989 68773: contig of 3886 bp in length
* 68774 68874: gap of 100 bp
* 68875 73467: contig of 4594 bp in length
* 73468 73567: gap of 100 bp
* 73568 77240: contig of 3673 bp in length
* 77241 77341: gap of 100 bp
* 77342 81090: contig of 3750 bp in length
* 81091 81190: gap of 100 bp
* 81191 85511: contig of 4321 bp in length
* 85512 85611: gap of 100 bp
```

```
* 85612 90487: contig of 4876 bp in length
* 90488 90587: gap of 100 bp
* 90588 95232: contig of 4705 bp in length
* 95233 95393: gap of 100 bp
* 95394 100087: contig of 4695 bp in length
* 100088 100187: gap of 100 bp
* 100188 106286: contig of 6099 bp in length
* 106287 106387: gap of 100 bp
* 106388 112851: contig of 6465 bp in length
* 112852 112951: gap of 100 bp
* 112952 117689: contig of 4738 bp in length
* 117690 117790: gap of 100 bp
* 117791 124437: contig of 6648 bp in length
* 124438 124537: gap of 100 bp
* 124538 133578: contig of 9041 bp in length
* 133579 133678: gap of 100 bp
* 133679 141606: contig of 7928 bp in length
* 141607 141706: gap of 100 bp
* 141707 154145: contig of 12439 bp in length
* 154146 154245: gap of 100 bp
* 154246 165714: contig of 11469 bp in length
* 165715 165815: gap of 100 bp
* 165816 183553: contig of 17739 bp in length.
FEATURES
      source
      1. .183553
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone_lib="RP11-28A1"
      /clone_lib="RP11-11 Human Male BAC"
      /note="assembly_fragment"
      1. .1318
      /note="assembly_fragment"
      1419. .2757
      /note="assembly_fragment"
      2858. .4221
      /note="assembly_fragment"
      4322. .5593
      /note="assembly_fragment"
      5694. .6731
      /note="assembly_fragment"
      6832. .8028
      /note="assembly_fragment"
      8129. .9540
      /note="assembly_fragment"
      9641. .11147
      /note="assembly_fragment"
      11248. .12587
      /note="assembly_fragment"
      12688. .13829
      /note="assembly_fragment"
      13930. .15248
      /note="assembly_fragment"
      15349. .16757
      /note="assembly_fragment"
      16858. .18403
      /note="assembly_fragment"
```

```
Query Match          29.8%; Score 615.6; DB 2; Length 183553;
Best Local Similarity 98.8%; Pred. No. 3e-98;
Matches 621; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GTGAATGTGAGGGTTTGTGACATTCAGATGCTTAGGAACACAGAGTGGGTGAGGGGGCCC 60
Db 103843 GTCAACGACAGGGGTTTCATGACATTCAGATGCTTAGGAACACAGAGTGGGTGAGGGGGCCC 103902

Qy 61 CAGCGAGGGCTGATTCCTTGGCGGAGAGAGTAGGTAAGGGTTTCGATGAGTCCCTT 120
Db 103903 CAGCGAGGGCTGATTCCTTGGCGGAGAGAGTAGGTAAGGGTTTCGATGAGTCCCTT 103962

Qy 121 AAAGGCAAAAGGTAACAGAGCCAGAGAGCTCGAGGGGAGACTTTCAGTTCAAGCCA 180
Db 103963 AAAGGCAAAAGGTAACAGAGCCAGAGAGCTCGAGGGGAGACTTTCAGTTCAAGCCA 104022

Qy 181 CAGATTGGTGGAAAGTGTGCGCGCCGCCGCTGCTCTCGAGCGGTGTCGACCTA 240
```


Db 121 TGTCTGGTGGTGGCTGCACTCGAATCCCGGCCACGACTCCAAAGGGCTCTCTCTCTCCCG 180
Qy 514 CTGGGATATCGGTCCGGCGCGCCAACTCCCAAGTTCGGCTCTCTCGGGGTCGGAGCACC 573
Db 181 CTGGGATATCGGTCCGGCGCGCCAACTCCCAAGTTCGGCTCTCTCGGGGTCGGAGCACC 240
Qy 574 AACACAGAGCCATCCGAGATGAGCAACAAGACGGGCATCATTTTACTTCGATCAGATCCTG 633
Db 241 AACACAGAGCCATCCGAGATGAGCAACAAGACGGGCATCATTTTACTTCGATCAGATCCTG 300
Qy 634 GTGAATGTGGTAAATTTTTCACATTCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 693
Db 301 GTGAATGTGGTAAATTTTTCACATTCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Qy 694 TACAGTTTCAGTTTTCAGTGAATTAAGTCTTACAGAGCCAAAATATCCAGGTTAACTTG 753
Db 361 TACAGTTTTCAGTTTTCAGTGAATTAAGTCTTACAGAGCCAAAATATCCAGGTTAACTTG 420
Qy 754 ATGTTAAATGGAACACCAAGTAATATCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 813
Db 421 ATGTTAAATGGAACACCAAGTAATATCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 480
Qy 814 GCTGCCACGAATGGTGTCT 873
Db 481 GCTGCCACGAATGGTGTCT 540
Qy 874 GAGAAAGGTAATTTGGTGGAGGCTGCGAGTATCCACGTTTCTGCTTCTGCTTCTGCTTCTGCT 933
Db 541 GAGAAAGGTAATTTGGTGGAGGCTGCGAGTATCCACGTTTCTGCTTCTGCTTCTGCTTCTGCT 600
Qy 934 CCCCTATAGGATTC 947
Db 601 CCCCTATAGGATTC 614

RESULT 11
AX498264
LOCUS AX498264 611 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 5 from Patent WO02055694.
ACCESSION AX498264
VERSION AX498264.1 GI:23343180
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Erickson,M.R., Bour,B.A., Bihain,B. and Tanaka,H.
Metabolic gene polymorphisms and polypeptides and uses thereof
Patent: WO 02055694-A 5 18-JUL-2002;
GENSET (FR)

FEATURES
source
1..611
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..606
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD48704.1"
/db_xref="GI:23343181"
/db_xref="REMBL:CAD48704"
/translation="MGSRRLASVPAVLVLTLPLGIPVVAQNDTEPILVLEKCLVVC
DSNPATDSKSSSSPLGISVRAANSKVAFAVSTNHEPSEMKNTRIIYFDQLVNV
GNFTLESVFAPKGIYSFVIVKYQSTQIVNMLNGKPVISAFAGDKDVTREA
ATNGVLLYLDKEDKVLKLEKGLVGVNQYSTFGFLVFLP"
ORIGIN

Query Match 29.5%; Score 609.4; DB 6; Length 611;
Best Local Similarity 99.8%; Pred. No. 1.5e-97;
Matches 610; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 337 ATGGGTCCCGGGCGCGGGCTGTCCGGTGCCTGGCGCGCTGTCTCTCTCTCTCTCTCTCTCT 396
Db 1 ATGGGTCCCGGGCGCGGGCTGTCCGGTGCCTGGCGCGCTGTCTCTCTCTCTCTCTCTCTCT 60
Qy 397 CGGGGCTGCCCTGTCTGGGCACAGAAACACACGAGGCCATTCGTCGTGGAGGGCAAGTGT 456
Db 61 CGGGGCTGCCCTGTCTGGGCACAGAAACACACGAGGCCATTCGTCGTGGAGGGCAAGTGT 120
Qy 457 CTGTGTGTGTGGTCTGAAACCGGGCCACGAGTCTCAAGGGCTCTCTCTCTCTCTCTCTCTCT 516
Db 121 CTGTGTGTGTGGTCTGAAACCGGGCCACGAGTCTCAAGGGCTCTCTCTCTCTCTCTCTCT 180
Qy 517 GGGATATCGGTTCGGGCGGCAACTCCAGGTGCGCTTCTCGCGGTCGCGAGCACCAAC 576
Db 181 GGGATATCGGTTCGGGCGGCAACTCCAGGTGCGCTTCTCGCGGTCGCGAGCACCAAC 240
Qy 577 CAGAGCCATCCGAGATGAGCAACAAGACGGGCATCATTTTACTTCGATCAGATCCTGCTG 636
Db 241 CAGAGCCATCCGAGATGAGCAACAAGACGGGCATCATTTTACTTCGATCAGATCCTGCTG 300
Qy 637 AATGTGGGTAAATTTTTCACATTCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 696
Db 301 AATGTGGGTAAATTTTTCACATTCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Qy 697 AGTTTCAGTTTTCAGTGAATTAAGTCTACAGAGCCAAAATATCCAGGTTAACTTGATG 756
Db 361 AGTTTCAGTTTTCAGTGAATTAAGTCTACAGAGCCAAAATATCCAGGTTAACTTGATG 420
Qy 757 TTAATGGAACACCAAGTAATATCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 816
Db 421 TTAATGGAACACCAAGTAATATCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 480
Qy 817 GCCACGAATGGTGTCT 876
Db 481 GCCACGAATGGTGTCT 540
Qy 877 AAGGTAATTTGGTGGAGGCTGCGAGTATCCACGTTTCTGCTTCTGCTTCTGCTTCTGCT 936
Db 541 AAGGTAATTTGGTGGAGGCTGCGAGTATCCACGTTTCTGCTTCTGCTTCTGCTTCTGCT 600
Qy 937 CTATAGGATTC 947
Db 601 CTATAGGATTC 611

RESULT 12
AL929097/c
LOCUS AL929097 107500 bp DNA linear ROD 06-SEP-2003
DEFINITION Mouse DNA sequence from clone RP23-76M8 on chromosome 2, complete
sequence.
ACCESSION AL929097
VERSION AL929097.12 GI:34495119
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 107500)
Almeida,J.
Direct Submission
Submitted (06-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On Sep 6, 2003 this sequence version replaced gi:33504039.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-76M8 is from the RPCI-23 Mouse BAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

Location/Qualifiers

1..107500

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="2"

/clone="RP23-76M8"

/clone_lib="RPCI-23"

ORIGIN

Query Match 29.4%; Score 608; DB 10; Length 107500;

Best Local Similarity 73.6%; Pred. No. 6.1e-97;

Matches 1009; Conservative 0; Mismatches 300; Indels 61; Gaps 16;

```

QY 737 CTATCAGGTTAACTGTGATGTTAAATGGAACACAGTAATATCTGCTTTGCCGGGGACA 796
Db 71306 CTTTCAGGTTAACTGTGATGTTAAATGGAACACAGTCATCTCTGCAATTTGCTGGTGATA 71247
QY 797 AGATGTTACTGTGAAGTCCGACGAAGTGTCTCTCTACCTAGATGATAAGAGGATA 856
Db 71246 AAGATGTGACCCGTGAAGCGCCACTAATGGAGTGTCTCTCTACCTGACGACAAAGAGATA 71187
QY 857 AGCTTTACCTAAACTGGAGAAAGTAAATTTGTTGGAGCTGGCAGTATTCACGTTTT 916
Db 71186 AGGTCTACCTAAACTGGAGAAAGTAAATTTGTTGGAGCTGGCAGTATTCACGTTTT 71127
QY 917 CTGGCTTTCTGTGTTCCCTATAGGATTCATTTCTCCATGATGTTATCCAGGTGAG 976
Db 71126 CTGGCTTTCTGTGTTCCCTATAGAACCCGACTTCTCGTGATGCCAAGAGTGAGG 71067
QY 977 GGATGCCCACTCCTGAGTATTTGGAAGATCATTTTTCATCATGATGATGTCCTTT 1036
Db 71066 CG-----CCACCCCGGTTATCAGAAGAGCACTCTTGGATGATGTCCTTACTGTTT 71012
QY 1037 ATTGGTTCTCATGGTGGATGATGATTTCTAAGGATTCAG-CTGCTGTAACCAATACA 1095
Db 71011 GTTCATGGC---TGACTATGGATTTCTCTCGAAGATTCAGACCTGCTGCGACAGACACA 70955
QY 1096 AAATTTTCACAGATTATT---TGTTGTGTCTGTTTCAGTATATTTGGATTGGAGCTCTA 1151
Db 70954 AGTTTCACAGATCAATTTGTTGTGTGTCGAGTTCTTATATGAGTGGAGCTGGACCTTA 70895
QY 1152 AGCAGATATPACC-----TAGCTTAAATGTAAACAGTCAAAAGCTGTCTCAAGACTTAT 1206
Db 70894 AAGCTGAAAAATCCCGATCTGCTTACATGTTAACTGTCAAGACCGCTCTCAAGTTGT 70835
QY 1207 TCTGAATTCATTCCTGGGATTTACTGAATAGTACAGATGTGGAA---TTTTATTGTT 1263
Db 70834 CCTG---TTTAGTTTCTGCTGTTACAGCATTTGGCTGCAGATGTGGGTTTTTTTATTCT 70777

```

```

QY 1264 TTAGTTTTAAAGACTGGCAACCCAGGCTTAAGGATTAGAAAACCTCTAAAGTTCTGACTTC 1323
Db 70776 TTGTTTTGAAGACAGGCTTAATGGGTCTGAAAGTTAGAAAACCTTGAAGTTCTGACTTC 70717
QY 1324 AATCAACGGTTAGTGTGATCTGCCAAGAAGACTGTATAC-----TGTC 1366
Db 70715 AGT----GGTATAGTGAAGTGTCCAAAGACATGTATACACGTTGATGATAT 70661
QY 1367 TTAATATATGATTAATTTGTTTATTTTATTTCCCTTGGAAATAGTTGTTGGTCTTGTATA 1426
Db 70660 TAATTATACCTTTTAAACAAAATATAATTAATTAATTTGTTGTTGTTGTTCTCTTAA 70601
QY 1427 AAAAATTTGATTTTTTTTTTCACTACTGTAATGTTTCTCTTAAATAAGTAATG 1486
Db 70600 GGTAACTGGATGTTATTTT-AGTGACTGGCTGTTGTTCTCTCTACCCCAAGGTAACA 70542
QY 1487 AATGGCTTCCCAACAATTTTACCTTGACTACGATATCATCGACATGACTTCTCT----CA 1542
Db 70541 AATGCTTCTGCTCAAAATGTCCTTAACCTGAACTCACCCACATGACCTCCCTCGCC 70482
QY 1543 AAAAATAAGATGCTTCATAGTTGTTTAAATGTTATGTAAGTGAAGAGATCATATTTCC 1602
Db 70481 CCTAAACAGATGCTCCATAGTTGTTTAAATTAATATATATGTAAGAGATCATATTTCC 70422
QY 1603 AGTTATATTTTCTTAAGAAGAGATAGATCATAAATCTGACAAAGAAAAGTTGCTTAC 1662
Db 70421 AAATTAATTTTCTTAAGAAAACAGATCAT-ATCTGACAAAGAGAGATTAATTTAC 70363
QY 1663 CCAAAATCTAAGTGCTCAATCCCTGAGCCTCGACAAACAGCTCCCTCGAGGGAAATC 1722
Db 70362 CCCAACTCTAA--TGCTCAGCGCCAAACCTTGCAAGTACCTCTCCTTGGTGGAAATC 70305
QY 1723 TTATACCTTATGCTCAACTTTTAAATTAAGTATGATTAATTAACCTTTTATTAATAAAC 1782
Db 70304 GAATACCTCTATCGCTCAACCTTTAAATTAACATGATGATTAATAGTACTACTTTTAAACCCA 70245
QY 1783 TAAGGTTTTTTTTTTCGTAGACATGACCACTTTTAACTGTTAACTGGTGGGATGCTGT 1842
Db 70244 TAAGG--AAATTTTCTCTCTCATGACCACTTTTAACTGGAGATGATGAGCCTT 70187
QY 1843 GTTCTTAATATACCTTTTCAAGCTTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1902
Db 70186 AATTTTAAATGATCTATTTTCAAATCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 70127
QY 1903 CCTTAACCTTTTAAATGTTATATTTATCTGTTTACCTTAATTAATTAATTAATTAATCC 1962
Db 70126 CCTTAATCTTAACCTTTGTTATATTTATCTGTTTACCTTAATTAATTAATTAATTAATCC 70067
QY 1963 ATATCTAAATTTAGTCAATATC-----TTGTTCTTTTGTATAGGTATGATTAATTAATTA 2017
Db 70065 ATAACTAAATTTAGTCAATATCTTTTGTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 70007
QY 2018 AATTAATTAAG--TCTGTTATGAATAAAGATTAATTAATTAATTAATTAATTAATTA 2065
Db 70006 AATTAATTAATGTTCTGTTTACAGATAAAGATTAATTAATTAATTAATTAATTAATTA 69957

```

RESULT 13
AL929254/c

LOCUS
DEFINITION

Mouse DNA sequence from clone Rp23-45611 on chromosome 2, complete sequence.

ACCESSION AL929254

VERSION AL929254.12

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 222104)

AUTHORS Leongamornlert, P.

TITLE Direct Submission

AL929254 222104 bp DNA linear ROD 15-SBP-2003

Mouse DNA sequence from clone Rp23-45611 on chromosome 2, complete sequence.

ACCESSION AL929254

VERSION AL929254.12

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 222104)

AUTHORS Leongamornlert, P.

TITLE Direct Submission

AC117551 256869 bp DNA linear HTG 03-MAR-2000
Mus musculus clone RP23-80N8, WORKING DRAFT SEQUENCE, 23 unordered
pieces.
AC117551
AC117551.5 GI:286311357
HTG; HTGS PHASE1; HTGS DRAFT.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Muridae; Murinae; Mus.
Birren,B., Nusbaum,C. and Lander,E.
1 (bases 1 to 256869)
Mus musculus, clone RP23-80N8
Unpublished
2 (bases 1 to 256869)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Bouckgaert,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,M., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kellis,C., LaRoque,K., Lamazares,R.,
Landers,T., Iehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Macdonald,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigliolo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 256869)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Bouckgaert,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,J., Faro,S.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Haggopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,M., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (03-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 3, 2003 this sequence version replaced gi|28570472.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

```
/clone="RP23-80N8"
/clone.lib="RP23-23 Female Mouse BAC"
1..5629
/feature="assembly_fragment"
vector_end:SP6
vector_side:left
5730..6454
/feature="assembly_fragment"
6555..7659
/feature="assembly_fragment"
7760..9340
/feature="assembly_fragment"
9441..11012
/feature="assembly_fragment"
11113..12927
/feature="assembly_fragment"
13028..14458
/feature="assembly_fragment"
14559..16483
/feature="assembly_fragment"
16584..18111
/feature="assembly_fragment"
18212..20336
/feature="assembly_fragment"
20437..23319
/feature="assembly_fragment"
23420..24924
/feature="assembly_fragment"
25025..28333
/feature="assembly_fragment"
28434..34747
/feature="assembly_fragment"
34848..42292
/feature="assembly_fragment"
42393..54917
/feature="assembly_fragment"
55018..74879
/feature="assembly_fragment"
74980..90451
/feature="assembly_fragment"
90552..108473
/feature="assembly_fragment"
108574..123030
/feature="assembly_fragment"
123131..177280
/feature="assembly_fragment"
177381..215343
/feature="assembly_fragment"
215444..256869
/feature="assembly_fragment"
clone_end:T7
vector_side:right

ORIGIN
Query Match 29.4%; Score 608; DB 2; Length 256869;
Best Local Similarity 73.6%; Pred. No. 7e-97;
Matches 1009; Conservative 0; Mismatches 300; Indels 61; Gaps 16;
QY 737 CTATCCAGGTAACTTGATGTTAAATGGAAACCAAGTAAATCTGCGCTTGGGGGACA 796
DB 247240 CTTTCCAGGTAACTTGATGTTAAATGGAAACCAAGTAAATCTGCGCTTGGGGGACA 247181
QY 797 AAGATGTTACTCGTGAAGTGCACCAAGTGGTCTGCTCTACCTAGATAAAGAGGATA 856
DB 247180 AAGATGTTACTCGTGAAGTGCACCAAGTGGTCTGCTCTACCTAGATAAAGAGGATA 247121
QY 857 AGGTTTACCTAAATCGGAGAAAGGTAATTTGGTGGAGGCTGGCAGTATTCACCGTTT 916
DB 247120 AGGTTTACCTAAATCGGAGAAAGGTAATTTGGTGGAGGCTGGCAGTATTCACCGTTT 247061
QY 917 CTGGCTTTCTGGTGTTCCTCTCTATAGGATTCATTTCTCATGATGTTCTATCCAGGTGAG 976
DB 247060 CTGGCTTTCTGGTGTTCCTCTCTATAGGATTCATTTCTCATGATGTTCTATCCAGGTGAG 247001
```

QY 2018 AATTATTATG--TCTGTATAGATAAAGATTAAATATATATGTTAAAAAA 2065
Db 245940 AATTATTATGCTCTGTGTACAGATAAAGATTATATATGTTAGTAAA 245891

RESULT 15
AC142108.2
WPCOMMENT
Sequence split into 4 fragments LOCUS AC142108 Accession AC142108
Fragment Name Begin End
AC142108.0 1 110000
AC142108.1 100001 210000
AC142108.2 200001 310000
AC142108.3 300001 377129
Continuation (3 of 4) of AC142108 from base 200001 (AC142108 Mus musculus chromosome UNK)

Query Match 28.8%; Score 594.6; DB 2; Length 110000;
Best Local Similarity 73.1%; Pred. No. 1.4e-94;
Matches 986; Conservative 0; Mismatches 304; Indels 59; Gaps 15;

QY 737 CTATCCAGGTTAACTTGATGTTAAATGGAACCAAGTAAATATCTGCTTGGGGGACA 796
Db 51523 CTTTCCAGGTTAACTTGATGTTAAATGGAACCAAGTAAATATCTGCTTGGGGTATA 51582

QY 797 AAGATGTTACTCTGTAAGCTGCCAGAAATGGTCTGCTCTACCTAGATAAAGAGATA 856
Db 51583 AAGATGTTACCCGTGAAGCGCCACTAATGGAGTCTCTGCTACCTGGACAAAGAGATA 51642

QY 857 AGGTTTACCTAAACTGGAGAAAGGTAATTTGGTGGAGGCTGGCAGTATTCACGTTTT 916
Db 51643 AGGTTTACCTAAACTGGAGAAAGGTAATTTGGTGGAGGCTGGCAGTATTCACGTTTT 51702

QY 917 CTGGCTTCTGGTGTTCCTCTATAGAACCCGACCTCTCGGTAGTCCCAAGAGTGAG 976
Db 51703 CTGGCTTCTGGTGTTCCTCTATAGAACCCGACCTCTCGGTAGTCCCAAGAGTGAGG 51762

QY 977 GGATGACCCACTCTCAGTATTATGGAAGATCAATTTTTCATCATGATGATGTTT 1036
Db 51763 CG-----CCACCCCGGTTATCAGAGAGCACTCTGGATGATGCTTACTGTTTT 51817

QY 1037 ATTGGTTTCTAGGGTGGATATGGAATTTCAAGATCTTAG-CCTGTCTGAACCAATACA 1095
Db 51818 GTTCATGGC---TGACTATGGAATCTCTCTGAGAATCTTAGACCTGTCTGGACAGACACA 51874

QY 1096 AATTTTACAGATATAT---TGTGTGTCTCTTTCAGTATATTTGGATGGGACTCTA 1151
Db 51875 AAGTTTACAGATATATTTCTGTGTGTGCGACGTTTCTTATATGTTGGAGCTGGACTCTA 51934

QY 1152 AGCAGATAATACC---TATGCTTAAATGTAAACAGTCAAAAGCTGTCTGCAAGACTTAT 1206
Db 51935 AAGCTGAATAATCCCGATCTGTTACATGTTAACTGTCAAGAGCCGTCTCAAGGTTGT 51994

QY 1207 TCTGAATTTCAATTTCTGGGATTTAGTAATTTAGTACAGATGTGGAA---TTTTATTGT 1263
Db 51995 CCTG---TTTATGTTTCTGCTGTACAGCAATGGCTGAGATGTGGGTTTTTTTATTCT 52052

QY 1264 TTAGTTTAAAGACTGGCAACCAAGGCTTAAGGATTTAGAAACTCTAAAGTTCTGACTTC 1323
Db 52053 TTGTTTGAAGACAGGCTTAATGGGCTGAAAGTTAGAAACCTTGAAGTTCTGACTTC 52112

QY 1324 AATCAACGGTTAGTGTGATAGTCCAAAGAACTGTATAC-----TGTTG 1366
Db 52113 AGT---GGTATAGTGGAACTGCAAGACATGTATACGTTGGATACATGATTAT 52168

QY 1367 TTAATAATTGATTAATTTGTTTATTTCTTTGGAAATTTAGTTTGGTTCTTGTGTA 1426
Db 52169 TAAATTATACITTTTAAACAAATTAATAATTTACTTTTGAATTTGTTTGTCTCTTAA 52228

QY 1427 AAAACTTGGATTTTCTTTCAGTAATGTTATTTCTTCTTAAATAAGGTAATG 1486
Db 52229 GGTAACTGGGATTTGATTTTT-AGTGACTGGGTTGTCTCTCTACCCATAAGGTAACA 52287

QY 1487 AATGGCTTGGCCCAAAATTTTACCTTTGACTAGATATCATCGACATGACTTCTCT-CA 1542

Search completed: May 15, 2004, 09:06:12
Job time : 8117 secs

Db 52288 AATGCTTTGGCTCAAAATGTCCTTTAACTCTGAACCTCACCACATGACCTCCTCCTGCC 52347
QY 1543 AAAAAAAGATGCTTCAATAGTTGATTTTAAATGATATGGAAGAGTCAATTTTCC 1602
Db 52348 CCTAAACGAATGCTCCATAGTTGATTTTAAATATATATGTAAGAGAGTCAATTTTCC 52407

QY 1603 AAGTTATATTTTCTAAGAAGAAATAGATCATATAATCTGCAAGGAAAAAGTTGCTTAC 1662
Db 52408 AAATTTATTTTCTAATAGGAAAAACAGATCAT-ATCTGCAAGAAGAAAGTTATTAC 52466

QY 1663 CCAAAATCTAAGTGTCTCAATCCTGAGCTCAGCAAAAACAGCTCCCTCCGAGGGAATC 1722
Db 52467 CCAAACTCTAA--TGCTCAGCGCCAAACCTTGGCAAGTACCTCTCTCTGTTGGGAAATC 52524

QY 1723 TTATACCTTTATGCTCAACTTTTAAATAAATGATTAATAACCACTTTATTAATAAAC 1782
Db 52525 GAATACCTCTATCGCTCAACCTTTAAATAACATGATTAATAAGTACTCTTTATTAACCA 52584

QY 1783 TAAGGTTTTTTTTTTTTCCTGAGACATGACCACTTTTATTAACCTGGTGGATGCTGT 1842
Db 52585 TAAGG--AATATTTTCTCCTTCACATGACCACTTTTAACTGGAGATAGTATGCCCTT 52642

QY 1843 GTTCTAATATPACCTATTTTTTCAAGGCTTCTGTGTATTTTGAAGTATCATCTGGTTTTG 1902
Db 52643 ATTTTAAATTTGATCTATTTTTCAAATCTCCTGTGTGTGTTGAAATATCATCTGGTTTTG 52702

QY 1903 CCTTAACCTTTTAAATTTGATATATTTATCTGTTAGCTAAATATAAATTCAAATATCC 1962
Db 52703 CCTTAATTCCTTAACCTTGATATATTTATCTGTTCAGCTAAATATAAATTCAGATACTCC 52762

QY 1963 ATATCTAAATTTAGTCAATATC-----TTGTCTTTTGTATAGGTCAATATGAATTCATAA 2017
Db 52763 CTAACTAAATTTAGTCAATATCTTATTTTGTCTTTTGTATAGGTCAATATGAATTCATAA 52822

QY 2018 AATTAATTTATGCTGTTATAGAAATAAGA 2046
Db 52823 AATTAATTTATGCTGTTATAGAAATAAGA 52851

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 04:14:15 ; Search time 5398 Seconds
(without alignments)
11423.750 Million cell updates/sec

Title: US-10-015-387A-219

Perfect score: 2065

Sequence: 1 gfgaatgtgaggggtttgatg.....attaatatgttaaaaaa 2065

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2065	100.0	3091	11	BC050026 Homo sapi
2	1077	52.2	2406	11	AK032406 Mus muscu
3	1012.8	49.0	2802	11	AK032621 Mus muscu
4	751	36.4	905	12	BI668562

5	686	33.2	696	13	EX093493
6	606	29.3	606	29	AY418431
7	603	28.2	606	29	AY418432
8	592.6	28.7	706	9	AV722278
9	572.4	27.7	579	13	EX506987
10	520.4	25.2	700	12	BI669085
11	504	24.4	504	9	AA868507
12	474	23.0	597	29	AY418433
13	454.8	22.0	745	14	CB318208
14	399.8	19.4	436	13	BU781883
15	396	19.2	750	13	BU382466
16	388.4	18.8	734	13	BU369123
17	388	18.8	421	9	AV721267
18	375.2	18.2	683	9	AV343573
19	357	17.3	597	9	AJ452689
20	334	16.2	340	9	AL732165
21	326.2	15.8	331	9	AI468259
22	323.6	15.7	566	10	BE755791
23	320.8	15.5	324	9	AW074518
24	315.6	15.3	674	9	AV348597
25	311.8	15.1	802	13	BQ444664
26	307.4	14.9	756	14	CE207272
27	292.4	14.2	1783	11	AK042555
28	292.4	14.2	2202	11	AK032112
29	290.4	14.1	754	13	BU140889
30	288	13.9	903	13	BO722274
31	287.6	13.9	346	9	AW047113
32	280	13.6	291	10	BF445600
33	275.4	13.3	689	13	BU457636
34	262.8	12.7	284	9	AL449740
35	259.8	12.6	513	13	BY372910
36	258.8	12.5	447	9	AI848962
37	253.4	12.3	1185	13	EX394263
38	247.8	12.0	907	14	CD254575
39	244.4	11.8	440	10	BB311702
40	240.4	11.6	380	13	BY630619
41	239.2	11.6	717	12	BM950922
42	238.2	11.5	635	10	BB622353
43	237.6	11.5	883	14	CD254586
44	235.4	11.4	422	10	BE779670
45	233.6	11.3	2428	11	AK076174

ALIGNMENTS

RESULT 1
BC050026

LOCUS

DEFINITION

Homo sapiens,

mRNA.

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

3091 bp mRNA linear HTC 10-APR-2003
cerebellin precursor-like 1, clone IMAGE:4820943,

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

3091 bp mRNA linear HTC 10-APR-2003
cerebellin precursor-like 1, clone IMAGE:4820943,

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

3091 bp mRNA linear HTC 10-APR-2003
cerebellin precursor-like 1, clone IMAGE:4820943,

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

BC050026

Sequencing Center

Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 71 Row: c Column: 9
 This clone has the following problem: retained intron.

FEATURES

source
 1..3091
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4820943"
 /tissue_type="Brain, hippocampus"
 /clone_lib="NIH MGC_95"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

ORIGIN

Query Match 100.0%; Score 2065; DB 11; Length 3091;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GTCAATGTGAGGTTTGATCAGCTTCAGATGCTTAGGAACAGAGTGGGTGCGAGGGCCC	60
DB	968	GTGAATGTGAGGTTTGATCAGCTTCAGATGCTTAGGAACAGAGTGGGTGCGAGGGCCC	1027
QY	61	CAGCGAGGGCTGATCTTGGCGGAGGAGAGTAGGTAAAGGGTCTGCGATGAGCTCCT	120
DB	1028	CAGCGAGGGCTGATCTTGGCGGAGGAGAGTAGGTAAAGGGTCTGCGATGAGCTCCT	1087
QY	121	AAAGGACAAAGGTAAACAGAGCCAGGAGAGAGCTCGAGGGGAGACTTTGACTCAAGCCA	180
DB	1088	AAAGGACAAAGGTAAACAGAGCCAGGAGAGAGCTCGAGGGGAGACTTTGACTCAAGCCA	1147
QY	181	CAGAAATGTGGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA	240
DB	1148	CAGAAATGTGGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA	1207
QY	241	GCGCTAGCATCTTCCGAGCACCGGATCCCGGGTAGGAGCGAGCGCGGGGAGCACC	300
DB	1208	GCGCTAGCATCTTCCGAGCACCGGATCCCGGGTAGGAGCGAGCGCGGGGAGCACC	1267
QY	301	AGCGCCAGCGGCTCGGCGTCCGACACAGGCTCACCATGGGCTCCGGGCGCGGGCGCTG	360
DB	1268	AGCGCCAGCGGCTCGGCGTCCGACACAGGCTCACCATGGGCTCCGGGCGCGGGCGCTG	1327
QY	361	TCGCGGTGCGGCGGTGCTGCTGCTCCTCAGCGTCCGCGGCTGCGCGGTGCGGCGACAG	420
DB	1328	TCGCGGTGCGGCGGTGCTGCTGCTCCTCAGCGTCCGCGGCTGCGCGGTGCGGCGACAG	1387
QY	421	AAGGACAGGAGCCCATCTGCTGCTGAGGCAAGTCTCTGCTGCTGCTGCTGCTGCTGCTG	480
DB	1388	AAGGACAGGAGCCCATCTGCTGCTGAGGCAAGTCTCTGCTGCTGCTGCTGCTGCTGCTG	1447
QY	481	GCCAGGACTCCAGGCTCCTCTCTCTCCCGCTGGGGATATCGGTCCGGGCGGCGCAAC	540
DB	1448	GCCAGGACTCCAGGCTCCTCTCTCTCCCGCTGGGGATATCGGTCCGGGCGGCGCAAC	1507
QY	541	TCCAGGTGCGCTTCTCGGCGGTGCGGAGCACCAACAGGCTCCGAGATGAGCAAC	600
DB	1508	TCCAGGTGCGCTTCTCGGCGGTGCGGAGCACCAACAGGCTCCGAGATGAGCAAC	1567
QY	601	AAGACGGCGATCATTTACTTCTGATCAGATCTCTGCTGATGCTGGGTAAATTTTTCACATG	660
DB	1568	AAGACGGCGATCATTTACTTCTGATCAGATCTCTGCTGATGCTGGGTAAATTTTTCACATG	1627
QY	661	GAGTCTGCTTTGTAGCACCAGAAAGGAATTTACAGTTTCAGTTTTCAGGTGATTA	720

DB	1628	GAGTCTGCTTTGTAGCACCAAGAAAGGAATTTACAGTTTCAGTTTTCAGGTGATTA	1687
QY	721	GTCTACAGAGCAAACTATCCAGGTAACTTGATGTTAAATGGAAGAAACAGTAATATCT	780
DB	1688	GTCTACAGAGCAAACTATCCAGGTAACTTGATGTTAAATGGAAGAAACAGTAATATCT	1747
QY	781	GCTTTTGGCGGACAAAGATGTTACTCTGTAAGCTGCCACGAATGCTCTCTGCTCTAC	840
DB	1748	GCTTTTGGCGGACAAAGATGTTACTCTGTAAGCTGCCACGAATGCTCTCTGCTCTAC	1807
QY	841	CTAGATAAAGAGGATAAGGTTTACCTAAACCTGGAGAAAGGTAATTTGGTGGAGGCTGG	900
DB	1808	CTAGATAAAGAGGATAAGGTTTACCTAAACCTGGAGAAAGGTAATTTGGTGGAGGCTGG	1867
QY	901	CAGTATTCACAGTTTCTGGCTTTCTGGTGTCTCCCTATAGGATTCATTTCTCCATGA	960
DB	1868	CAGTATTCACAGTTTCTGGCTTTCTGGTGTCTCCCTATAGGATTCATTTCTCCATGA	1927
QY	961	TGTTTATCCAGGTGAGGATGACCCACTCTGAGTTTATTTGGAAGATCATTTTTCATCAT	1020
DB	1928	TGTTTATCCAGGTGAGGATGACCCACTCTGAGTTTATTTGGAAGATCATTTTTCATCAT	1987
QY	1021	TGGATTGATGCTTTTATTTGGTTTCTCATGGTGGATATGGATTCTAAGGATTCAGCCT	1080
DB	1988	TGGATTGATGCTTTTATTTGGTTTCTCATGGTGGATATGGATTCTAAGGATTCAGCCT	2047
QY	1081	GTCTGAACCAATACAAAATTTCCAGATTAATTTGTTGTTGTTGTTTTCAGTATATTGGA	1140
DB	2048	GTCTGAACCAATACAAAATTTCCAGATTAATTTGTTGTTGTTGTTTTCAGTATATTGGA	2107
QY	1141	TTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTAACAGTCAAAAGCTGTCTCAAG	1200
DB	2108	TTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTAACAGTCAAAAGCTGTCTCAAG	2167
QY	1201	ACTTATCTGAATTTCAATTTCTGGGATTACTGAATTAAGTTCAGATGTTGAATTTAT	1260
DB	2168	ACTTATCTGAATTTCAATTTCTGGGATTACTGAATTAAGTTCAGATGTTGAATTTAT	2227
QY	1261	TGTTTATGTTTAAAGACTGGCAACCCAGCTCTAAGGATTAAGAACTCTAAGGTTCTGAC	1320
DB	2228	TGTTTATGTTTAAAGACTGGCAACCCAGCTCTAAGGATTAAGAACTCTAAGGTTCTGAC	2287
QY	1321	TTCAATCAACGGTGTAGTGTATCTGCCAAAGACTGTATCTGTCTTAAATATATTGAT	1380
DB	2288	TTCAATCAACGGTGTAGTGTATCTGCCAAAGACTGTATCTGTCTTAAATATATTGAT	2347
QY	1381	ATATTTGTTTATTTTCTTTTGGAAATTAAGTTTGTGTTGTTGTTTCTGTAAGAACTTTG	1440
DB	2348	ATATTTGTTTATTTTCTTTTGGAAATTAAGTTTGTGTTGTTGTTTCTGTAAGAACTTTG	2407
QY	1441	TTTTTTTTCAGTAACTGGTATTAATGTTTCTTTAAATAAGGTAATGAATGGCTTGCCAC	1500
DB	2408	TTTTTTTTCAGTAACTGGTATTAATGTTTCTTTAAATAAGGTAATGAATGGCTTGCCAC	2467
QY	1501	AAATTTACCTTGATACGATATCATCGACATGCTTCTCAAAAAGAAAGAAAGTCTTCA	1560
DB	2468	AAATTTACCTTGATACGATATCATCGACATGCTTCTCAAAAAGAAAGAAAGTCTTCA	2527
QY	1561	TAGTGTGATTTTAAATTTGATATGTGAAGAGTCATATTTTCCAGTTATATTTTCTAAGA	1620
DB	2528	TAGTGTGATTTTAAATTTGATATGTGAAGAGTCATATTTTCCAGTTATATTTTCTAAGA	2587
QY	1621	AGAAGAAATAGATCATAAATCTGACAGGAAAGGTTGCTTACCAGAAATCTAAGTCTCA	1680
DB	2588	AGAAGAAATAGATCATAAATCTGACAGGAAAGGTTGCTTACCAGAAATCTAAGTCTCA	2647
QY	1681	ATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGAAATCTTATATCTTTATGCTCAA	1740
DB	2648	ATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGAAATCTTATATCTTTATGCTCAA	2707
QY	1741	CTTTAATTAATGATGATTAATAACACCTTTATTAATAAACCTTAAGGTTTCTTTTTC	1800

```

Db      2708 CTTTAATAAATGATGATAATAAACCACTTTTATTAATAAACCACTTAAGGTTTTTTTTTTTC 2767
QY      1801 CGTAGACATGACACATTTTATTAACCTGCTGGTGGATGCTGTTCTTTCTAATATATACCTAT 1860
Db      2768 CGTAGACATGACACATTTTATTAACCTGCTGGTGGATGCTGTTCTTTCTAATATATACCTAT 2827
QY      1861 TTTTCAAGGCTTCTGTTGTTATTTGAAGTATCATCTGTTTGGCTTAACTCTTTAAATTTG 1920
Db      2828 TTTTCAAGGCTTCTGTTGTTATTTGAAGTATCATCTGTTTGGCTTAACTCTTTAAATTTG 2887
QY      1921 TATATATTTATCTGTTAGCTAATATTAATTCATATCCCATATCTAAATTTAGTGA 1980
Db      2888 TATATATTTATCTGTTAGCTAATATTAATTCATATCCCATATCTAAATTTAGTGA 2947
QY      1981 ATATCTGTTCTTTGTTAGCTATATGAATTCATATAAATTTATTTATGCTGTTATAGAA 2040
Db      2948 ATATCTGTTCTTTGTTAGCTATATGAATTCATATAAATTTATTTATGCTGTTATAGAA 3007
QY      2041 TAAAGATTAATATATGTTAAAAAA 2065
Db      3008 TAAAGATTAATATATGTTAAAAAA 3032

RESULT 2
AK032406
LOCUS   2406 bp mRNA linear HTC 18-SEP-2003
DEFINITION Mus musculus adult male olfactory brain cDNA, RIKEN full-length
            GLYCOPROTEIN PRECURSOR, full insert sequence.
ACCESSION AK032406
VERSION   1 GI:26328228
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
          Carninci, P. and Hayashizaki, Y.
          High-efficiency full-length cDNA cloning
          Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
REFERENCE 2
          Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
          Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
          Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new genes
          Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
REFERENCE 3
          Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P.,
          Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M.,
          Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
          Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
          Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
          Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
          Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
          RIKEN integrated sequence analysis (RISA) system--384-format
          sequencing pipeline with 384 multicapillary sequencer
          Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
REFERENCE 4
          The RIKEN Genome Exploration Research Group Phase II Team and the
          PANTOM Consortium.
          Functional annotation of a full-length mouse cDNA collection
          Nature 409, 685-690 (2001)
5
          The PANTOM Consortium and the RIKEN Genome Exploration Research
          Group Phase I & II Team.
          Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs

```

JOURNAL
REFERENCE
AUTHORS

Nature 420, 563-573 (2002)
6 (bases 1 to 2406)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

TITLE

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source

Location/Qualifiers
1..2406
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6430538G21"
/db_xref="MGI:2395903"
/db_xref="taxon:10090"
/clone="6430538G21"
/sex="male"
/tissue_type="olfactory brain"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
673..1269
/notes="unnamed protein product; CEREBELLIN-LIKE
GLYCOPROTEIN PRECURSOR (SWISSPROT|Q9NTU7, evidence: PASTY,
96%ID, 100%length, match=594)
putative"
/codon_start=1
/protein_id="BAC27855.1"
/db_xref="GI:26328229"

CDS

translation="MGSARRALSVVPAVLILLVPAQNDTEPIVLEKGLVVCDSN
PANDSGSSSSPLGISVRANSKVPASVSTNHPSEMSNKTITVFDQILVNVGNF
FILESVFVATRKGIYSFHVIVYQSQIQLVNLNGKPVISAFAGDKDVTREATN
GVLLYLDKDKVYKLEKNGLLGGWQVSTFSGFLVFLP"
2387..2392
/note="putative"
2406
/note="putative"

polyA_signal

polyA_site

ORIGIN

Query Match 52.2%; Score 1077; DB 11; Length 2406;
Best Local Similarity 77.0%; Pred. No. 6.5e-183;
Matches 1624; Conservative 0; Mismatches 395; Indels 90; Gaps 22;

QY 1 GTCAATGTAGGGCTTTGATGACTTT-CAGATGCTAGGACCAAGTGGTGGAGGGGCC 59

Db 337 GTCACTGAGAGGGCTCTGGAGACTTTTAGATGCAAGGGGACAAAGTGGGCGCGTGGCC 336

QY 60 CCAGGCGAGGCGCTGATTCTTGGCGCGAGGAGTAGGTAAGGGTTCATGATGAGTCTCT 119

Db 397 CCAGGCGAAGAACTGATTGGGGGGTGGAG--GAGTTAGGGTAGGAGTTCATGATGAGTCTCT 454

241	AGAAGAAGATAGATCATTAATCTGACAGAGGAAAAAGTTGCTTACCCAAAAATCTAAGTGC	300
1678	TCAATCCCTGAGCCTCAGCAAAAACAGCTCCCTCCGAGGGGAAATCTTATATCTTATTGCT	1737
301	TCAATCCCTGAGCCTCAGCAAAAACAGCTCCCTCCGAGGGGAAATCTTATATCTTATTGCT	360
1738	CAACTTTTAATTAAAAATGATTGATTAATACACCTTTATTAACAACTTAAGGTTTTTTTTTT	1797
361	CAACTTTTAATTAAAAATGATTGATTAATACACCTTTATTAACAACTTAAGGTTTTTTTTTT	420
1798	TTCCGTAGACATGACCACTTTATTAACTGGTGGGATGCTGTGTTTCTTAATTATPACC	1857
421	TTCCGTAGACATGACCACTTTATTAACTGGTGGGATGCTGTGTTTCTTAATTATPACC	480
1858	TATTTTTCAAGGCTTCTGTTGTATTTGAAAGTATCATCTGTTTTTGCCTTAACTCTTAAA	1917
481	TATTTTTCAAGGCTTCTGTTGTATTTGAAAGTATCATCTGTTTTTGCCTTAACTCTTAAA	540
1918	TTGTATATATTTATCTGTTTAGCTAAATATTAAATTTCAAATATCCCATATCTAAATTAGT	1977
541	TTGTATATATTTATCTGTTTAGCTAAATATTAAATTTCAAATATCCCATATCTAAATTAGT	600
1978	GCAATATCTTGTCTTTTGTATAGCTCATATGAATTCATAAAATPATTTTATGTCGTGTATA	2037
601	GCAATATCTTGTCTTTTGTATAGCTCATATGAATTCATRAAATPATTTTATGTCGTGTATA	660
2038	GAATAAAGATTAAATATATGTTAAAAAAA	2065
661	GAATAAAGATTAAATATATGTTAAAAAAA	688

RESULT 6	606 bp	DNA	linear	GSS 17-DEC-2003
AY418431				
LOCUS	Homo sapiens CBLN1L1 gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	genomic survey sequence.			
ACCESSION	AY418431			
VERSION	AY418431.1	GI:39774391		
KEYWORDS	GSS.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 606)			
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,			
	Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,			
	Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,			
	Adams,M.D. and Cargill,M.			
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous			
	gene trios			
JOURNAL	Science 302 (5652), 1960-1963 (2003)			
PUBMED	14671302			
REFERENCE	2 (bases 1 to 606)			
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,			
	Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,			
	Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,			
	Adams,M.D. and Cargill,M.			
TITLE	Direct Submission			
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,			
	Rockville, MD 20850, USA			
COMMENT	These sequences were made by sequencing genomic exons and ordering			
	them based on alignment			
FEATURES	Location/Qualifiers			
source	1..606			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
	<1..>606			
gene	/gene="CBLN1L1"			
	/locus_tag="HCM6555"			
ORIGIN				
Query Match	29.3%	Score 606;	DB 29;	Length 606;

Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT These sequences were made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 source
 1..606
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..5606
 /gene="CELNLI"
 /locus_tag="HCM655"

ORIGIN
 Query Match 29.2%; Score 603; DB 29; Length 606;
 Best Local Similarity 99.5%; Pred. No. 6.6e-98;
 Matches 603; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 337 ATGGGCTCCGGGCGCGGGCGCTGTCTCGCGGTGCGGGCGCTGTCTGTGTCTTCCAGCTG 396
 |||||
 Db 1 ATGGGCTCCGGGCGCGGGCGCTGTCTCGCGGTGCGGGCGCTGTCTGTGTCTTCCAGCTG 60
 |||||

QY 397 CCGGGCTGCCCTCTGGGACACAGACACACGAGCCCATCTGTGTGGAGGCAAGTGT 456
 |||||
 Db 61 CCGGGCTGCCCTCTGGGACACAGACACACGAGCCCATCTGTGTGGAGGCAAGTGT 120
 |||||

QY 457 CTGGTGTGTGGACTCGAACCCGGCCAGGACTCCAAAGGCTCTTCTTCTCCCGCTG 516
 |||||
 Db 121 CTGGTGTGTGGACTCGAACCCGGCCAGGACTCCAAAGGCTCTTCTTCTCCCGCTG 180
 |||||

QY 517 GGGATATCGGTCCGGGCGGCAACTCCAAAGTGTCTTCTCGCGGTGCGGGACCAAC 576
 |||||
 Db 181 GGGATATCGGTCCGGGCGGCAACTCCAAAGTGTCTTCTCGCGGTGCGGGACCAAC 240
 |||||

QY 577 CAGAGCCATCCAGATAGCAACAGACCGCATCATTTACTTCGATCAGATCTCTGGTG 636
 |||||
 Db 241 CAGAGCCATCCAGATAGCAACAGACCGCATCATTTACTTCGATCAGATCTCTGGTG 300
 |||||

QY 637 AATGTGGTAAATTTTTCATTTGAGTGTCTTTGTAGTACCAAGGCAAGAAAGAAATTC 696
 |||||
 Db 301 AATGTGGTAAATTTTTCATTTGAGTGTCTTTGTAGTACCAAGGCAAGAAAGAAATTC 360
 |||||

QY 697 AGTTTCAGTTTTCAGTGATTAAGTCTACAGAGCCAACTATCCAGGTTCATTTGATG 756
 |||||
 Db 361 AGTTTCAGTTTTCAGTGATTAAGTCTACAGAGCCAACTATCCAGGTTCATTTGATG 420
 |||||

QY 757 TTAATGGAACCCAGTAATATCTGCTTTGCGGGGCAAGAGATGTTACTCGTGAAGCT 816
 |||||
 Db 421 TTAATGGAACCCAGTAATATCTGCTTTGCGGGGCAAGAGATGTTACTCGTGAAGCT 480
 |||||

QY 817 GCCACGAATGGTCTCTGCTCTACCTAGATAAAGAGATAAGTTCCTTAAACCTGGAG 876
 |||||
 Db 481 GCCACGAATGGTCTCTGCTCTACCTAGATAAAGAGATAAGTTCCTTAAACCTGGAG 540
 |||||

QY 877 AAAGTAATTTGGTGGAGCTGGAGTATTCACGTTTCTGGCTTCTGTGTTTCCCT 936
 |||||
 Db 541 AAAGTAATTTGGTGGAGCTGGAGTATTCACGTTTCTGGCTTCTGTGTTTCCCT 600
 |||||

QY 937 CTATAG 942
 |||||
 Db 601 CTATAG 606
 |||||

RESULT 8
 AV722278
 LOCUS AV722278 HTB Homo sapiens cDNA clone HTBZF01 5', mRNA linear EST 16-OCT-2000
 DEFINITION AV722278
 ACCESSION AV722278
 VERSION AV722278.1 GI:10824606
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 706)

AUTHORS

Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X.,
 Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L.,
 Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R.,
 Chen, J., Chen, Z. and Han, Z.

TITLE

Homo sapiens cDNA HTB clones

JOURNAL

Unpublished (2000)

COMMENT

Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

FEATURES

source

1..706

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="HTBZF01"

/tissue_type="Hypothalamus"

/dev_stage="Adult"

/lab_host="SOLR"

/clone_lib="HTB"

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:

XhoI"

ORIGIN

Query Match 28.7%; Score 592.6; DB 9; Length 706;
 Best Local Similarity 96.3%; Pred. No. 4.8e-96;
 Matches 629; Conservative 0; Mismatches 16; Indels 8; Gaps 2;

QY 526 GTCCGGCGGCCCAACTCCAAAGGTGCGCTTCTCGCGGTGCGGAGCACCAACACGAGCCA 585
 Db 1 GTCCGGTGGCCCAACTCCCAATGTGCGCTTCTCGCGGTGCGGAGCACCAACACGAGCCA 60
 |||||

QY 586 TCCGAGATGAGCAACAGACGGGCATCATTTACTTCGATCAGATCTCTGGTGAATGGGT 645
 |||||
 Db 61 TCCGAGATGAGCAACAGACGGGCATCATTTACTTCGATCAGATCTCTGGTGAATGGGT 120
 |||||

QY 646 AATTTTTCACATTTGAGTCTCTTTGTAGCACCAAGAAAGGAATTTACAGTTTCAGT 705
 |||||
 Db 121 AATTTTTCACATTTGAGTCTCTTTGTAGCACCAAGAAAGGAATTTACAGTTTCAGT 180
 |||||

QY 706 TTTCAGTGATTAAGTCTTACAGAGCCAACTATCCAGTTTAACTTGAATGGA 765
 |||||
 Db 181 TTTCAGTGATTAAGTCTTACAGAGCCAACTATCCAGTTTAACTTGAATGGA 240
 |||||

QY 766 AAACCAAGTAATATCTGCTTTTGGGGGCAAGATGTTACTCTGTAAGTCCAGGAAT 825
 |||||
 Db 241 AAACCAAGTAATATCTGCTTTTGGGGGCAAGATGTTACTCTGTAAGTCCAGGAAT 300
 |||||

QY 826 GGTGTCTCTCTCTACCTAGATAAAGAGATAAGTTTACTTAAACTCGAGAAAGTAAT 885
 |||||
 Db 301 GGTGTCTCTCTCTACCTAGATAAAGAGATAAGTTTACTTAAACTCGAGAAAGTAAT 360
 |||||

QY 886 TTGGTTGGAGCTGGCAGTATTCACGTTTCTCGGCTTCTCGGTGTCCTCCCTATAGAT 945
 |||||
 Db 361 TTGGTTGGAGCTGGCAGTATTCACGTTTCTCGGCTTCTCGGTGTCCTCCCTATAGAT 420
 |||||

QY 946 TCAATTTCTCCATGATGTTTCATCCAGGTGAGGATGACCCACTCTGAGTTATGGAAGA 1005
 |||||
 Db 421 TCAATTTCTCCATGATGTTTCATCCAGGTGAGGATGACCCACTCTGAGTTATGGAAGA 480
 |||||

QY 1006 TCATTTTTCATCATTTGATTCATGTCCTTTTATTTGTTTCTCATCGGTGATGATTC 1065
 |||||
 Db 481 TCATTTTTCATCATTTGATTCATGTCCTTTTATTTGTTTCTCATCGGTGATGATTC 540
 |||||

QY 1066 TAAGGATTCAGCTGTCTGAAACCAATACAAAATTTACAGATTATTTGTGTGTCTGT 1125
 |||||


```

RESULT 11
AA868507/c
OCUS
DEFINITION
ak43b06.s1 Soares testis_NHT Homo sapiens cdna clone IMAGE:1408691
3', mRNA sequence.
ACCESSION
AA868507
KEYWORDS
AA868507.1 GI:2963952
EST.
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 504)
REFERENCE
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgaps-remail.nih.gov
cdna library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cdna library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLN at:
www-bio1.llnl.gov/bbrp/image/image.html
Insert Length: 667 Std Error: 0.00

```

RESULT	12
AY418433	
LOCUS	
DEFINITION	AY418433 Mus musculus CBLNLI gene, partial sequence, genomic survey sequence.
ACCESSION	AY418433
VERSION	AY418433.1 GI:39774393
KEYWORDS	GSS.

AUTHORS

Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,
Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392

COMMENT

Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1. 750
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST86801"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN75"
/note="Organ: trunks; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Query March 19.2%; Score 396; DB 13; Length 750;
Best Local Similarity 76.5%; Pred. No. 8.8e-61;
Matches 578; Conservative 0; Mismatches 165; Indels 13; Gaps 7;
QY 572 CCAACACGAGCCATCGAGATGAGCAACAGACGCGCATCATTTACTTCGATCAGATCC 631
Db 1 CCAACACGAGCCCTCGAGATGAGCAACAGACGCGCATCATTTACTTCGACAGATCC 60
QY 632 TGGTGAATGTGGGTAATTTTTCACATTTGGAGTCTGCTTTGTAGACACCAAGAAAGGAA 691
Db 61 TAGTAAACGTGGGCAATTTTTCACGTTGGAACTCTGCTTTGTCCACCAAGAAAGGAA 120
QY 692 TTTCAGGTTTCAGTTTTCAGTGAATTAAGTCTACAGAGCCAAACTATCCAGGTTAACT 751
Db 121 TTTCAGGTTTCAGTTTTCAGTGAATTAAGTCTACAGAGCCAAACTATCCAGGTTAACT 180
QY 752 TGAATGTTAAATGGAAACAGTAATATCTGCTTTTCGGGGGACAAAGATGTTACTCGTG 811
Db 181 TGATGCTCAATGGAAAGCCAGTCATCTGCTTTTGTGGGACAAAGGACGTCACCTCGTG 240
QY 812 AAGCTGCCACGAGATGCTGCTCTACCTAGATTAAGAGGATAGGTTTACCTAAAC 871
Db 241 AAGCTGCCACGAGATGCTGCTCTATCTAGCAAGGAGGATAAGGTTTACCTGAAGC 300
QY 872 TGGAGAAAGGTAATTTGGTTGGAGGCTGGCAGTATTCACGTTTCTCGCTTTCTGGCT 931
Db 301 TGGAGAAAGGTAATCTGGTTGGTGGATGGCAGTATTCAGTTTCTGGCTTTCTGGCT 360
QY 932 TCCCCCTATAGGATTCATTTCTCCATGATGTTATCCAGGTGAGGATGACCCACTCCT 991
Db 361 TTCCCCCTGTA-AAGTCAATTTTCTGATGATGTTGCTCCAGTGTGGAGTCGTCATTGCT 419
QY 992 GAGTTATTGGAAGATCATTTTTCATCTCATTTGGATTGATGTCCT-TTATTGGTTTCTCATG 1050

Db 420 CTGTCACATGAAGATCATCTCTGTCATCGTGGGATTTGATGTTTCTTATTTGTTTTCATG 479
QY 1051 GGTGGATATGGATTC---TAAGGATTTCTAGCCTGTCTGAAACCAATACAAATTTTCACAG 1106
Db 480 GGTGAATATAGATTTCTCTGTATGGATTTTGGCCCCATCTGAACCTACACAGAGTTTCACA 539
QY 1107 ATATTTGTGTGTCTGTCTTTTCAGTATATTTGGATTGGGACTCTA---AGCAGATAATACC 1164
Db 540 GAATTTCTGTGTGT-TTAAATACAAATATTTGGATTGAGACTAAAGCAGACGATAATAA 598
QY 1165 TATGCTTTAAATGTAACAGTCAAAAGCTGTCTGCAAGACTTATTTCTGAATTTCTTCTG 1224
Db 599 CTGTGCTTAATGTTTACGGTCTTAAAGCTGCTGCACGGTTTATTCAG-GTATCGTATCTG 657
QY 1225 GGATTTACTGAATTTAGTTTACAGATGTGGAATTTTATTTTGTGTTTAAAGACTGGCAA 1284
Db 658 GGACTTTATGATTCATGAGAGAGTGA---TTTCCTTTTAACTGTGAAACGACTGGCAA 714
QY 1285 CCAGTCTTAAGGATTTAGAAACTCTAAAGTTCTGAC 1320
Db 715 CCAGGTCTATATTAATTAAGCCAGAAATTTGAGTTCAGAC 750

Search completed: May 15, 2004, 06:48:05

Job time : 5405 secs